Step by Step Guide for SPSS

IBM SPSS Statistics 28 Brief Guide

SPSS Software

အခြေခံ နည်းလမ်းများ

Dr Nay Win Aung

ဒေါက်တာနေဝင်းအောင် (ဖွဲ့ဖြိုးမှုပညာ)

မာတိကာ

Chapter ၁. Introduction (မိတ်ဆက်) ၁

Sample Files ၁

Opening a Data File (Data File တစ်ခုကို ဖွင့်ခြင်း) ၁

Running an Analysis (Analysis တစ်ခုကို Run ခြင်း) ၃

Creating Charts (Charts များ ဖန်တီးခြင်း) ၅

Chapter ၂. Reading Data (Data များကို ဖတ်ယူခြင်း) ၈

Basic Structure of IBM SPSS Statistics Data Files (IBM SPSS Statistics Data Files များရဲ့ အခြေခံ ဖွဲ့စည်းပုံ) ၈

Reading IBM SPSS Statistics Data Files (IBM SPSS Statistics Data Files များကို ဖတ်ယူခြင်း) ၈

Reading Excel Data (Excel Data များကို ဖတ်ယူခြင်း) ၉

Reading Data from a Database (Database မှ Data များကို ဖတ်ယူခြင်း) ၁၁

Reading Data from a Text File (Text File မှ Data များကို ဖတ်ယူခြင်း) ၁၅

Chapter ၃. Using the Data Editor ၁၉

Entering Numeric Data ၁၉

Entering String Data ၂၁

Defining Data ၂၂

Adding Variable Labels ၂၂

Changing Variable Type and Format ၂၃

Adding Value Labels ၂၄

Handling Missing Data ၂၅

Missing Values for a Numeric Variable ၂၆

Missing Values for a String Variable ၂၇

Chapter ၄. Examining Summary Statistics for Individual Variables ၂၈

Level of Measurement ၂၈

Summary Measures for Categorical Data ၂၉
Charts for Categorical Data ၃၀
Summary Measures for Scale Variables ၃၁
Histograms for Scale Variables ၃၃
Chapter ၅. Creating and editing charts ၃၄
Chart creation basics ၃၄
Using the Chart Builder gallery ၃၅
Defining variables and statistics ၃၆
Adding text ၃၈
Creating the chart ၃၉
Chapter ၆. Working with Output ၄၀
Using the Viewer ၄၀
Using the Pivot Table Editor ၄၂
Accessing Output Definitions ၄၂
Pivoting Tables ၄၂
Creating and Displaying Layers ၄၄
Editing Tables ၄၅
Hiding Rows and Columns ၄၆
Changing Data Display Formats ၄၆
TableLooks ၄၈
Using Predefined Formats ၄၈
Customizing TableLook Styles ၄၉
Changing the Default Table Formats ၅၂
Customizing the Initial Display Settings ၅၃
Displaying Variable and Value Labels ၅၄
Using Results in Other Applications ၅၆

Pasting Results as Word Tables ၅၆
Pasting Results as Text ၅၆
Exporting Results to Microsoft Word, PowerPoint, and Excel Files ၅၇
Exporting Results to PDF ၆၃
Exporting Results to HTML ၆၅
Chapter ၇. Working with Syntax ၆၇
Pasting Syntax ၆၇
Editing Syntax ၆၈
Opening and Running a Syntax File ၇၀
Using Breakpoints ၇၀
Chapter ၈. Modifying Data Values ၇၂
Creating a Categorical Variable from a Scale Variable ၇၂
Computing New Variables ၇၆
Using Functions in Expressions ၇၇
Pasting a Function into an Expression ၇၈
Editing a Function in an Expression ၇၉
Using Conditional Expressions ၇၉
Working with Dates and Times ၈၀
Calculating the Length of Time between Two Dates ၈၂
Adding a Duration to a Date ၈၃
Chapter ၉. Sorting and Selecting Data ၈၅
Sorting Data ၈၅
Split-File Processing ၈၆
Sorting Cases for Split-File Processing ၈၈
Turning Split-File Processing On and Off ၈၈
Selecting Subsets of Cases ၈၉

Selecting Cases Based on Conditional Expressions ၉၀
Selecting a Random Sample ၉၁
Selecting a Time Range or Case Range ၉၂
Treatment of Unselected Cases ၉၄
Case Selection Status ၉၅
Chapter ၁၀. Sample Files (နမူနာ ဖိုင်များ) ၉၆

Chapter 1. Introduction (မိတ်ဆက်)

ဒီ guide ကနေ ရရှိနိုင်တဲ့ features တော်တော်များများကို ဘယ်လိုအသုံးပြုရမလဲဆိုတာ ပြသပေးသွားမှာပါ။ အဆင့်လိုက် လက်တွေ့လုပ်ဆောင်နိုင်တဲ့ guide တစ်ခုအနေနဲ့ ဒီဇိုင်းထုတ်ထားတာဖြစ်ပါတယ်။ ဥပမာတွေမှာပြထားတဲ့ files တွေအားလုံးကို application နဲ့အတူ install လုပ်ထားတဲ့အတွက်ကြောင့်၊ ဒီမှာပြထားတဲ့ analyses တွေကို ကိုယ်တိုင်လိုက်လုပ်ကြည့်ပြီး တူညီတဲ့ results တွေကို ရရှိနိုင်မှာဖြစ်ပါတယ်။ တကယ်လို့ statistical analysis techniques အမျိုးမျိုးရဲ့ အသေးစိတ်ဥပမာတွေကို သိချင်တယ်ဆိုရင်တော့ Help menu မှာရနိုင်တဲ့ အဆင့်လိုက် Case Studies တွေကို စမ်းကြည့်နိုင်ပါတယ်။

Sample Files

ဒီမှာပြထားတဲ့ ဥပမာအများစုက demo.sav ဆိုတဲ့ data file ကို အသုံးပြုထားတာပါ။ ဒီ data file ဟာ လူထောင်ပေါင်းများစွာဆီကနေ ကောက်ယူထားတဲ့ ဖန်တီးထားတဲ့ survey တစ်ခုဖြစ်ပြီး၊ အခြေခံ demographic နဲ့ consumer သတင်းအချက်အလက်တွေ ပါဝင်ပါတယ်။ တကယ်လို့ သင်က Student version ကို သုံးနေတာဆိုရင်တော့၊ သင့်ရဲ့ demo.sav က မူရင်း data file ရဲ့ ကိုယ်စားပြု sample တစ်ခုသာဖြစ်ပြီး 1,500-case limit နဲ့ကိုက်ညီအောင် လျှော့ချထားတာဖြစ်ပါတယ်။ အဲဒီ data file ကိုသုံးပြီး သင်ရလာမယ့် results တွေက ဒီမှာပြထားတဲ့ results တွေနဲ့ ကွာခြားနေမှာပါ။

Product နဲ့အတူ install လုပ်ထားတဲ့ sample files တွေကို installation directory ရဲ့ Samples subdirectory ထဲမှာ ရှာတွေ့နိုင်ပါတယ်။ အင်္ဂလိပ်၊ ပြင်သစ်၊ ဂျာမန်၊ အီတလီ၊ ဂျပန်၊ ကိုရီးယား၊ ပိုလန်၊ ရုရှား၊ ရိုးရှင်းသောတရုတ်၊ စပိန်နဲ့ ရိုးရာတရုတ်ဘာသာစကား တစ်ခုချင်းစီအတွက် Samples subdirectory ထဲမှာ သီးခြား folder တစ်ခုစီ ရှိပါတယ်။ Sample files အားလုံးကို ဘာသာစကားအားလုံးနဲ့ မရနိုင်ပါဘူး။ ဘာသာစကားတစ်ခုမှာ sample file တစ်ခု မရှိဘူးဆိုရင်၊ အဲဒီ ဘာသာစကား folder ထဲမှာ အင်္ဂလိပ် version နဲ့ sample file ကို ထည့်သွင်းပေးထားပါတယ်။

Opening a Data File (Data File တစ်ခုကို ဖွင့်ခြင်း)

Data file တစ်ခုကို ဖွင့်ဖို့အတွက်-

၁။ menus တွေထဲကနေ File > Open > Data... ကို ရွေးချယ်ပါ။

ဖိုင်တွေဖွင့်ဖို့အတွက် dialog box တစ်ခု ပေါ်လာပါလိမ့်မယ်။ ပုံမှန်အားဖြင့် (By default) IBM SPSS Statistics data files တွေ (.sav extension) ကို ပြသပေးထားပါတယ်။ ဒီဥပမာမှာ demo.sav ဖိုင်ကို သုံးထားပါတယ်။

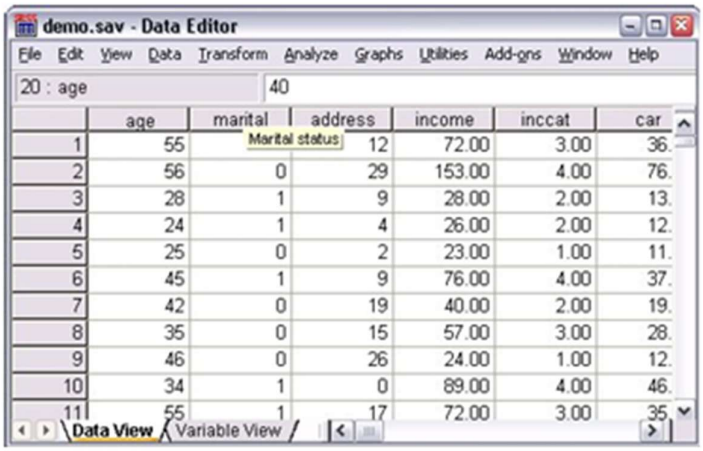


Figure 1. demo.sav file in Data Editor

Data file ကို Data Editor ထဲမှာ ပြသပေးပါလိမ့်မယ်။

Data View မှာ၊ တကယ်လို့ သင်က mouse cursor ကို variable name (column headings တွေ) ပေါ်မှာ တင်လိုက်မယ်ဆိုရင် ပိုပြီးအသေးစိတ်ကျတဲ့ variable label ကို ပြသပေးမှာပါ (အဲဒီ variable အတွက် label သတ်မှတ်ထားခဲ့ရင်ပေါ့) ။ပုံမှန်အားဖြင့်တော့ အမှန်တကယ် data values တွေကိုပဲ ပြသပေးပါတယ်။ labels တွေကို ပြချင်တယ်ဆိုရင်တော့-

၂။ menus တွေထဲကနေ View > Value Labels ကိုရွေးပါ။

Figure 2. Value Labels button

နောက်တစ်နည်းအနေနဲ့ toolbar ပေါ်က Value Labels button ကိုလည်း သုံးနိုင်ပါတယ်။

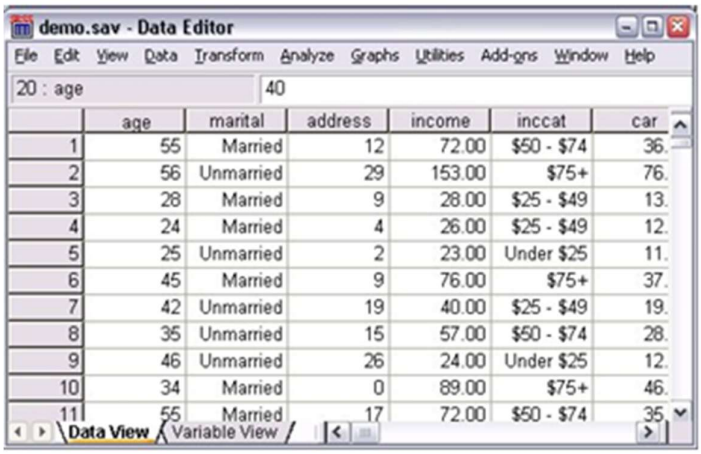

demo.sav - Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
20 : age
40
age marital address income inccat car
1 55 Married 12 72.00 $50 - $74 36.
2 56 Unmarried 29 153.00 $75+ 76.
3 28 Married 9 28.00 $25 - $49 13.
4 24 Married 4 26.00 $25 - $49 12.
5 25 Unmarried 2 23.00 Under $25 11.
6 45 Married 9 76.00 $75+ 37.
7 42 Unmarried 19 40.00 $25 - $49 19.
8 35 Unmarried 15 57.00 $50 - $74 28.
9 46 Unmarried 26 24.00 Under $25 12.
10 34 Married 0 89.00 $75+ 46.
11 55 Married 17 72.00 $50 - $74 35
Data View
Variable View

Figure 3. Value labels displayed in the Data Editor

အခုဆိုရင် အဖြေတွေကို အလွယ်တကူ နားလည်နိုင်ဖို့အတွက် descriptive value labels တွေကို ပြသပေးထားပြီ ဖြစ်ပါတယ်။

Running an Analysis (Analysis တစ်ခုကို Run ခြင်း)

တကယ်လို့ သင့်ဆီမှာ add-on options တွေရှိတယ်ဆိုရင် Analyze menu ထဲမှာ reporting နဲ့ statistical analysis categories စာရင်းတွေ ပါဝင်ပါတယ်။ ရိုးရှင်းတဲ့ frequency table (table of counts) တစ်ခုကို စတင်ပြီး ဖန်တီးကြည့်ကြပါမယ်။ ဒီဥပမာအတွက် Statistics Base option လိုအပ်ပါတယ်။

၁။ menus ထဲကနေ Analyze > Descriptive Statistics > Frequencies... ကိုရွေးပါ။

Frequencies dialog box ပေါ်လာပါလိမ့်မယ်။

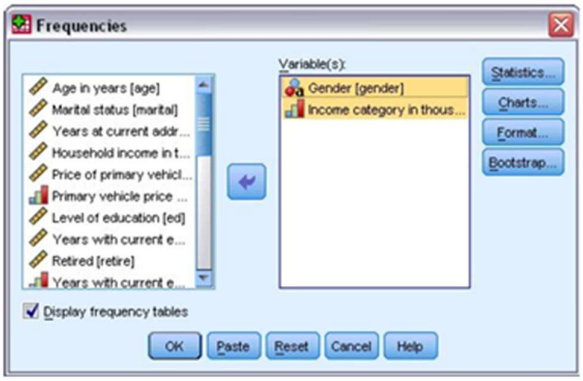


Figure 4. Frequencies dialog box

variable တစ်ခုစီရဲ့ဘေးက icon လေးက data type နဲ့ level of measurement အကြောင်း သတင်းအချက်အလက်ကို ပေးပါတယ်။

	Numeric	String	Date	Time
Scale (Continuous)		n/a		
Ordinal				
Nominal				

တကယ်လို့ စာရင်းထဲမှာ variable label ဒါမှမဟုတ် name က ပြတ်နေတယ်ဆိုရင်၊ ၎င်းပေါ်ကို cursor တင်လိုက်တဲ့အခါ အပြည့်အစုံ label/name ကို ပြပေးပါလိမ့်မယ်။ variable name ဖြစ်တဲ့ inccat ကို အသေးစိတ် variable label ရဲ့နောက်က square brackets လေးတွေထဲမှာ

ပြသပေးထားပါတယ်။ Income category in thousands ဆိုတာက variable label ပါ။ တကယ်လို့သာ variable label မရှိခဲ့ရင်၊ list box ထဲမှာ variable name သီးသန့်ပဲ ပေါ်နေမှာပါ။

ပြင်ပ ဘောင်တွေ ဒါမှမဟုတ် ထောင့်တွေကို ကလစ်နှိပ်ပြီး ဆွဲချဲ့ခြင်းဖြင့် dialog boxes တွေကို windows တွေလိုပဲ အရွယ်အစား ပြောင်းလဲနိုင်ပါတယ်။ ဥပမာ၊ သင်က dialog box ကို ပိုကျယ်အောင်လုပ်လိုက်ရင် variable lists တွေလည်း ပိုကျယ်လာမှာပါ။

dialog box ထဲမှာ သင် analyze လုပ်ချင်တဲ့ variables တွေကို ဘယ်ဘက်က source list ကနေရွေးချယ်ပြီး ညာဘက်က Variable(s) list ထဲကို ဆွဲထည့် (drag and drop) နိုင်ပါတယ်။ Analysis ကို run ပေးမယ့် OK button က Variable(s) list ထဲကို အနည်းဆုံး variable တစ်ခု မထည့်မချင်း ပိတ်ထား (disabled) မှာဖြစ်ပါတယ်။ dialogs အများစုမှာ list ထဲက ဘယ် variable name ကိုမဆို right-click နှိပ်ပြီး pop-up menu ကနေ **Variable Information** ကိုရွေးချယ်ခြင်းဖြင့် နောက်ထပ် သတင်းအချက်အလက်တွေကို ရယူနိုင်ပါတယ်။

၂။ source variable list ထဲက Gender [gender] ကို ကလစ်နှိပ်ပြီး၊ အဲဒီ variable ကို ပစ်မှတ် Variable(s) list ထဲကို ဆွဲထည့်ပါ။

၃။ source list ထဲက Income category in thousands [inccat] ကို ကလစ်နှိပ်ပြီး၊ target list ထဲကို ဆွဲထည့်ပါ။

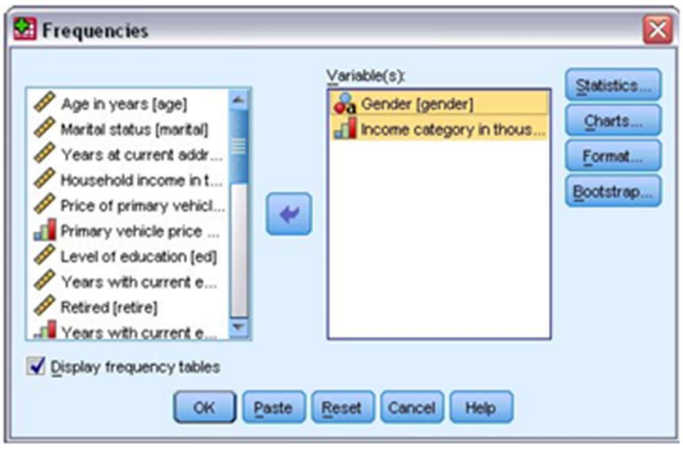


Figure 5. Variables selected for analysis

၄။ procedure ကို run ဖို့အတွက် OK ကို နှိပ်ပါ။ Results တွေကို Viewer window ထဲမှာ ပြသပေးပါလိမ့်မယ်။

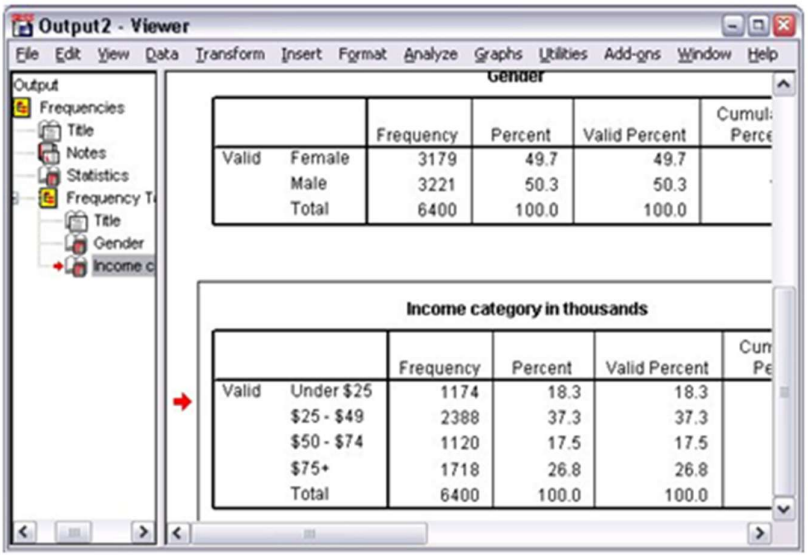


Figure 6. Frequency table of income categories

Creating Charts (Charts များ ဖန်တီးခြင်း)

တချို့ statistical procedures တွေက charts တွေကို ဖန်တီးနိုင်ပေမယ့်၊ charts တွေဖန်တီးဖို့အတွက် Graphs menu ကိုလည်း အသုံးပြုနိုင်ပါတယ်။ ဥပမာအားဖြင့် wireless telephone service နဲ့ PDA (personal digital assistant) ပိုင်ဆိုင်မှုကြားက ဆက်စပ်မှုကို ပြသတဲ့ chart တစ်ခုကို ဖန်တီးနိုင်ပါတယ်။

၁။ menus ထဲကနေ Graphs > Chart Builder... ကို ရွေးချယ်ပါ။

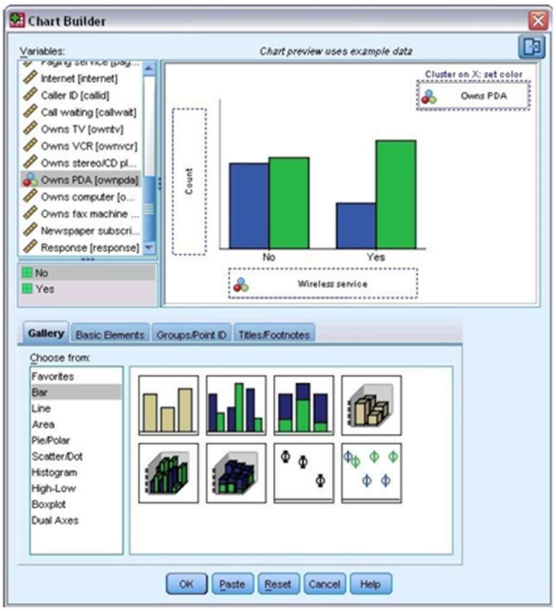


Figure 7. Chart Builder dialog box with completed drop zones

၂။ Gallery tab ကို နှိပ်ပါ (ရွေးချယ်ထားခြင်းမရှိသေးရင်ပေါ့)။

၃။ Bar ကို နှိပ်ပါ (ရွေးချယ်ထားခြင်းမရှိသေးရင်ပေါ့)။

၄။ Clustered Bar icon ကို Gallery အပေါ်က ဧရိယာအကျယ်ကြီးဖြစ်တဲ့ canvas ပေါ်ကို ဆွဲတင်ပါ။

၅။ Variables list ကို အောက်ဆွဲချပြီး Wireless service [wireless] ကို right-click နှိပ်ပါ၊ ပြီးရင် သူ့ရဲ့ measurement level အဖြစ် Nominal ကို ရွေးချယ်ပါ။

၆။ Wireless service [wireless] variable ကို x axis ဆီ ဆွဲထည့်ပါ။

၇။ Owns PDA [ownpda] ကို right-click နှိပ်ပြီး သူ့ရဲ့ measurement level အဖြစ် Nominal ကို ရွေးချယ်ပါ။

၈။ Owns PDA [ownpda] variable ကို canvas ရဲ့ ညာဘက်အပေါ်ထောင့်မှာရှိတဲ့ cluster drop zone ဆီ ဆွဲထည့်ပါ။

၉။ chart ကို ဖန်တီးဖို့ OK ကို နှိပ်ပါ။

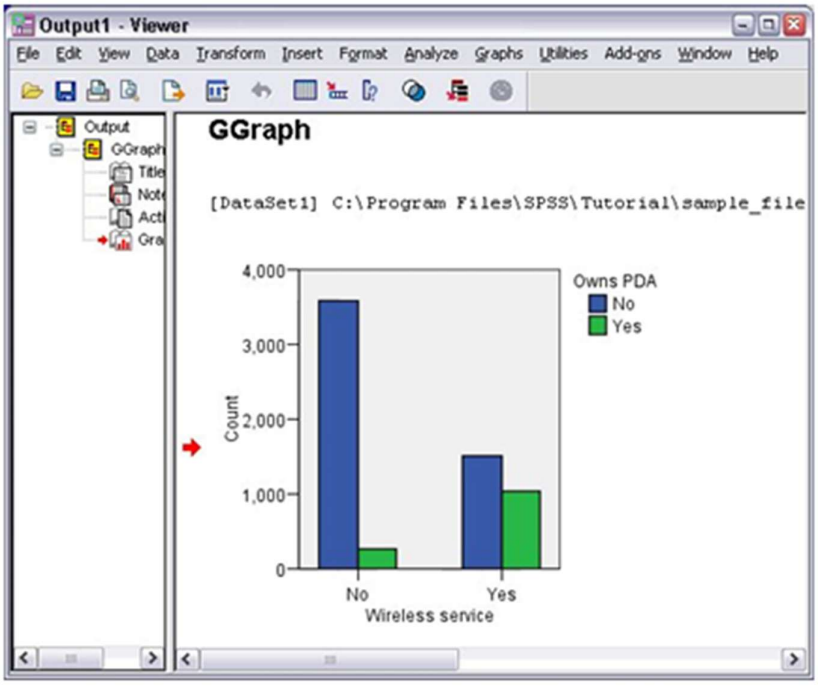


Figure 8. Bar chart displayed in Viewer window

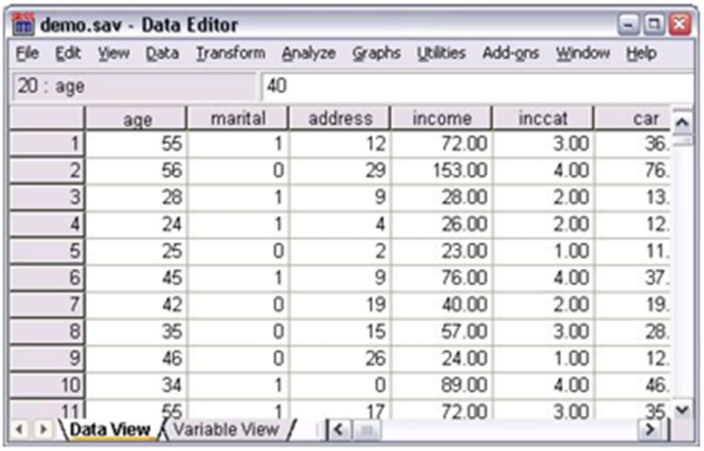
Bar chart ကို Viewer ထဲမှာ ပြသပေးပါလိမ့်မယ်။ ဒီ chart က wireless phone service သုံးတဲ့သူတွေဟာ wireless service မရှိတဲ့သူတွေထက် PDAs ပိုင်ဆိုင်ဖို့ အများကြီး ပိုများတယ်ဆိုတာကို ပြသနေပါတယ်။ Viewer window ရဲ့ contents pane ထဲမှာ charts တွေနဲ့

tables တွေကို double-click နှိပ်ပြီး edit လုပ်နိုင်ပါတယ်၊ ပြီးတော့ သင့်ရဲ့ results တွေကို တခြား applications တွေဆီကိုလည်း copy, paste လုပ်နိုင်ပါတယ်။ ဒီအကြောင်းအရာတွေကို နောက်ပိုင်းမှာ ဆက်လက် ရှင်းပြပေးသွားမှာပါ။

Chapter 2. Reading Data (Data များကို ဖတ်ယူခြင်း)

Data တွေကို တိုက်ရိုက် ရိုက်ထည့်လို့ရသလို၊ တခြား sources တွေကနေလည်း import လုပ်ယူလို့ရပါတယ်။ ဒီ chapter မှာတော့ IBM SPSS Statistics data files တွေ၊ Microsoft Excel လို spreadsheet application တွေ၊ Microsoft Access လို database application တွေနဲ့ text files တွေထဲမှာ သိမ်းထားတဲ့ data တွေကို ဘယ်လို read လုပ်မလဲ (ဖတ်ယူမလဲ) ဆိုတဲ့ လုပ်ငန်းစဉ်တွေကို ဆွေးနွေးသွားမှာပါ။

Basic Structure of IBM SPSS Statistics Data Files (IBM SPSS Statistics Data Files များရဲ့ အခြေခံ ဖွဲ့စည်းပုံ)



	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Figure 9. Data Editor

IBM SPSS Statistics data files တွေကို cases (rows တွေ) နဲ့ variables (columns တွေ) အနေနဲ့ ဖွဲ့စည်းထားပါတယ်။ ဒီ data file မှာဆိုရင် cases တွေက survey ဖြေဆိုသူ တစ်ဦးချင်းစီကို ကိုယ်စားပြုပါတယ်။ Variables တွေကတော့ survey မှာမေးထားတဲ့ မေးခွန်းတစ်ခုချင်းစီအတွက် အဖြေတွေကို ကိုယ်စားပြုတာပါ။

Reading IBM SPSS Statistics Data Files (IBM SPSS Statistics Data Files များကို ဖတ်ယူခြင်း)

.sav file extension ရှိတဲ့ IBM SPSS Statistics data files တွေထဲမှာ သင်သိမ်းဆည်းထားတဲ့ data တွေ ပါဝင်ပါတယ်။

၁။ menus တွေထဲကနေ File > Open > Data... ကို ရွေးချယ်ပါ။

၂။ demo.sav ဆိုတဲ့ဖိုင်ကို ရှာပြီး ဖွင့်လိုက်ပါ (ပိုသိချင်ရင် Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ ကြည့်နိုင်ပါတယ်)။

အခုဆိုရင် Data Editor ထဲမှာ data တွေကို မြင်ရပြီဖြစ်ပါတယ်။

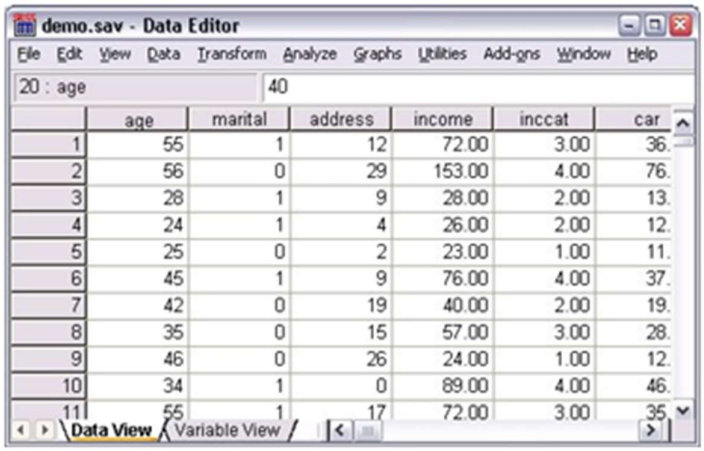


Figure 10. Opened data file

Reading Excel Data (Excel Data များကို ဖတ်ယူခြင်း)

Data အားလုံးကို Data Editor ထဲ တိုက်ရိုက်ရိုက်ထည့်မယ့်အစား၊ Microsoft Excel လို applications တွေကနေ data တွေကို ဖတ်ယူလို့ရပါတယ်။ Column headings တွေကိုလည်း variable names တွေအနေနဲ့ ဖတ်ယူနိုင်ပါသေးတယ်။

၁။ menus တွေထဲကနေ File > Import Data > Excel ကို ရွေးချယ်ပါ။

၂။ Samples\English folder ထဲကိုသွားပြီး demo.xlsx ကို ရွေးပါ။

Read Excel File dialog မှာ data file ရဲ့ preview ကို ပြသပေးပါလိမ့်မယ်။ ဖိုင်ထဲက ပထမဆုံး sheet ရဲ့ အကြောင်းအရာတွေကို ပြသပေးထားတာပါ။ တကယ်လို့ ဖိုင်ထဲမှာ sheets တွေ အများကြီးရှိနေရင် list ထဲကနေ လိုချင်တဲ့ sheet ကို ရွေးချယ်နိုင်ပါတယ်။ Gender အတွက် string values တချို့ရဲ့ ရှေ့မှာ spaces (ကွက်လပ်တွေ) ပါနေတာကို တွေ့ရပါလိမ့်မယ်။ MaritalStatus အတွက် values တချို့ကိုတော့ periods (.) အနေနဲ့ ပြသထားပါတယ်။

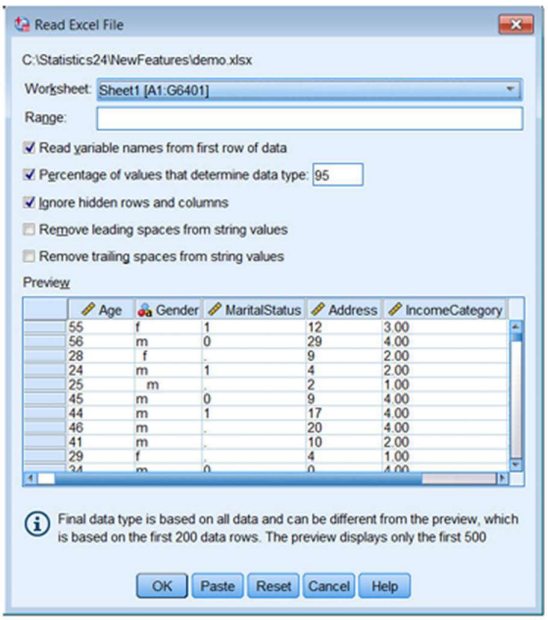
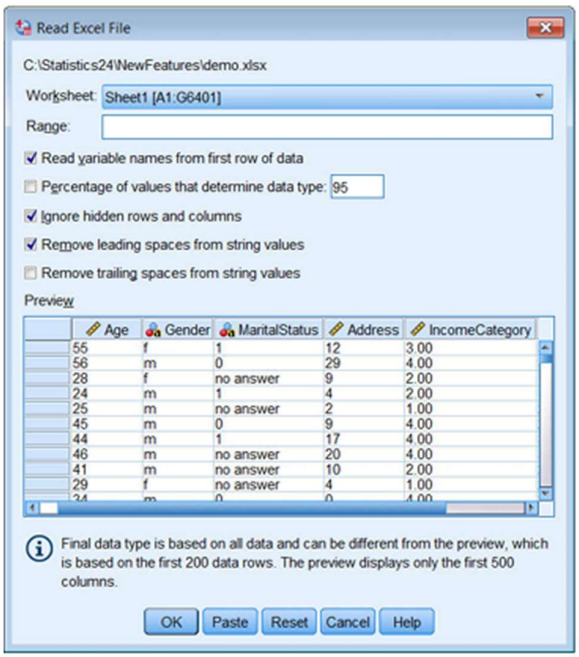

Read Excel File
C:\Statistics24\NewFeatures\demo.xlsx
Worksheet: Sheet1 [A1:G6401]
Range:
Read variable names from first row of data
Percentage of values that determine data type: 95
Ignore hidden rows and columns
Remove leading spaces from string values
Remove trailing spaces from string values
Preview
Age Gender MaritalStatus Address IncomeCategory
55 f 1 12 3.00
56 m 0 29 4.00
28 f . 9 2.00
24 m 1 4 2.00
25 m . 2 1.00
45 m 0 9 4.00
44 m 1 17 4.00
46 m . 20 4.00
41 m . 10 2.00
29 f . 4 1.00
Final data type is based on all data and can be different from the preview, which is based on the first 200 data rows. The preview displays only the first 500
OK Paste Reset Cancel Help

Figure 11. Read Excel File dialog

၃။ Read variable names from the first row of data ဆိုတာကို ရွေးချယ်ထားဖို့ သေချာပါစေ။ တကယ်လို့ column headings တွေက variable name rules တွေနဲ့ မကိုက်ညီဘူးဆိုရင်၊ မှန်ကန်တဲ့ variable names တွေအဖြစ် ပြောင်းလဲပေးသွားမှာပါ။ မူလ column headings တွေကိုတော့ variable labels အနေနဲ့ သိမ်းဆည်းပေးပါတယ်။

၄။ Remove leading spaces from string values ကို ရွေးချယ်ပါ။

၅။ Percentage of values that determine data type ကို အမှန်ခြစ် ဖြုတ်လိုက်ပါ။


Read Excel File
C:\Statistics24\NewFeatures\demo.xlsx
Worksheet: Sheet1 [A1:G6401]
Range:
Read variable names from first row of data
Percentage of values that determine data type: 95
Ignore hidden rows and columns
Remove leading spaces from string values
Remove trailing spaces from string values
Preview
Age Gender MaritalStatus Address IncomeCategory
55 f 1 12 3.00
56 m 0 29 4.00
28 f no answer 9 2.00
24 m 1 4 2.00
25 m no answer 2 1.00
45 m 0 9 4.00
44 m 1 17 4.00
46 m no answer 20 4.00
41 m no answer 10 2.00
29 f no answer 4 1.00
Final data type is based on all data and can be different from the preview, which is based on the first 200 data rows. The preview displays only the first 500 columns.
OK Paste Reset Cancel Help

အခုဆိုရင် system-missing ဖြစ်နေတဲ့ cells တွေမှာ "no answer" ဆိုတဲ့ string value ကို ပြသပေးနေပြီဖြစ်ပါတယ်။ တကယ်လို့ percentage of values parameter မရှိဘဲ column ထဲမှာ data type မျိုးစုံ ရောနှောနေမယ်ဆိုရင် အဲဒီ variable ကို string data type အနေနဲ့ ဖတ်ယူမှာပါ။ Values အားလုံးကို မူလအတိုင်း ထိန်းသိမ်းထားပေးပေမယ့် numeric values တွေကိုလည်း string values တွေလိုပဲ သတ်မှတ်သွားမှာဖြစ်ပါတယ်။

၆။ MaritalStatus ကို numeric variable တစ်ခုအနေနဲ့ သတ်မှတ်ဖို့အတွက် Percentage of values that determine data type ကို အမှန်ခြစ် (select) လုပ်ပေးပါ။

၇။ Excel file ကို ဖတ်ဖို့အတွက် OK ကိုနှိပ်ပါ။

အခုဆိုရင် data တွေကို Data Editor ထဲမှာ မြင်ရမှာဖြစ်ပြီး၊ column headings တွေကို variable names တွေအဖြစ် အသုံးပြုထားမှာပါ။ Variable names တွေမှာ spaces ပါလို့မရတဲ့အတွက် မူလ column headings တွေမှာပါတဲ့ spaces တွေကိုတော့ ဖယ်ရှားပေးသွားမှာပါ။ ဥပမာ၊ "Marital Status" ဆိုတဲ့ column heading ဟာ MaritalStatus ဆိုတဲ့ variable အဖြစ် ပြောင်းလဲသွားမှာပါ။ မူလ column heading ကိုတော့ variable label အနေနဲ့ ဆက်လက်ထိန်းသိမ်းထားမှာ ဖြစ်ပါတယ်။

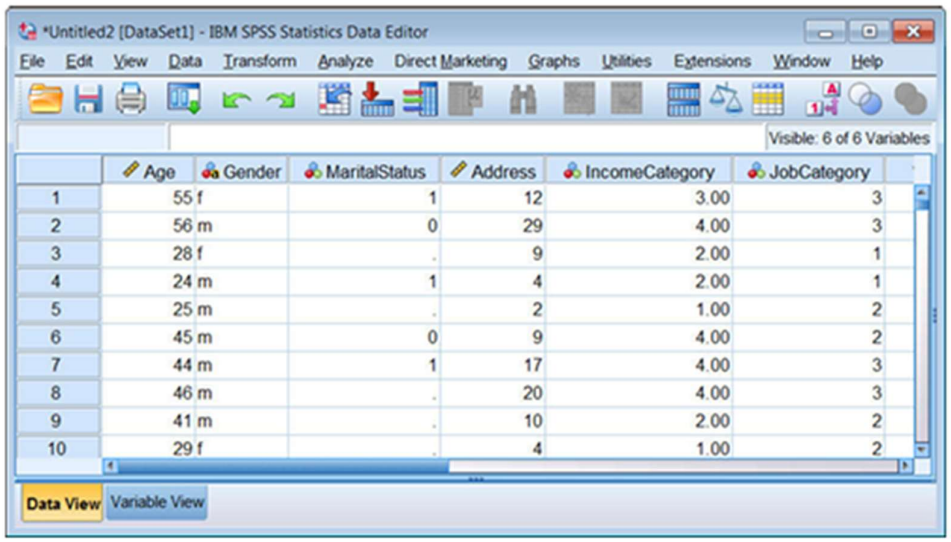

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help
Visible: 6 of 6 Variables
Age Gender MaritalStatus Address IncomeCategory JobCategory
1 55 f 1 12 3.00 3
2 56 m 0 29 4.00 3
3 28 f . 9 2.00 1
4 24 m 1 4 2.00 1
5 25 m . 2 1.00 2
6 45 m 0 9 4.00 2
7 44 m 1 17 4.00 3
8 46 m . 20 4.00 3
9 41 m . 10 2.00 2
10 29 f . 4 1.00 2
Data View
Variable View

Figure 12. Imported Excel data

Reading Data from a Database (Database မှ Data များကို ဖတ်ယူခြင်း)

Database sources တွေက data တွေကို Database Wizard သုံးပြီး အလွယ်တကူ import လုပ်ယူနိုင်ပါတယ်။ ODBC (Open Database Connectivity) drivers တွေကို သုံးတဲ့ ဘယ် database ကိုမဆို drivers တွေ install လုပ်ပြီးတာနဲ့ တိုက်ရိုက်ဖတ်ယူနိုင်ပါတယ်။ Database formats

တော်တော်များများအတွက် ODBC drivers တွေကို installation CD ထဲမှာ ထည့်သွင်းပေးထားပါတယ်။ တခြား drivers တွေကိုတော့ third-party vendors တွေဆီကနေ ရယူနိုင်ပါတယ်။ အသုံးအများဆုံး database application တစ်ခုဖြစ်တဲ့ Microsoft Access အကြောင်းကို ဒီဥပမာမှာ ဆွေးနွေးသွားမှာပါ။ မှတ်ချက် - ဒီဥပမာက Microsoft Windows အတွက် သီးသန့်ဖြစ်ပြီး Access အတွက် ODBC driver လိုအပ်ပါတယ်။ Microsoft Access ODBC driver က 32-bit version ရှိတဲ့ IBM SPSS Statistics နဲ့ပဲ အလုပ်လုပ်တာပါ။ တခြား platforms တွေမှာလည်း အဆင့်တွေက ဆင်တူပေမယ့် Access အတွက် third-party ODBC driver တော့ လိုအပ်နိုင်ပါတယ်။

၁။ menus တွေထဲကနေ File > Import Data > Database > New Query... ကို ရွေးချယ်ပါ။

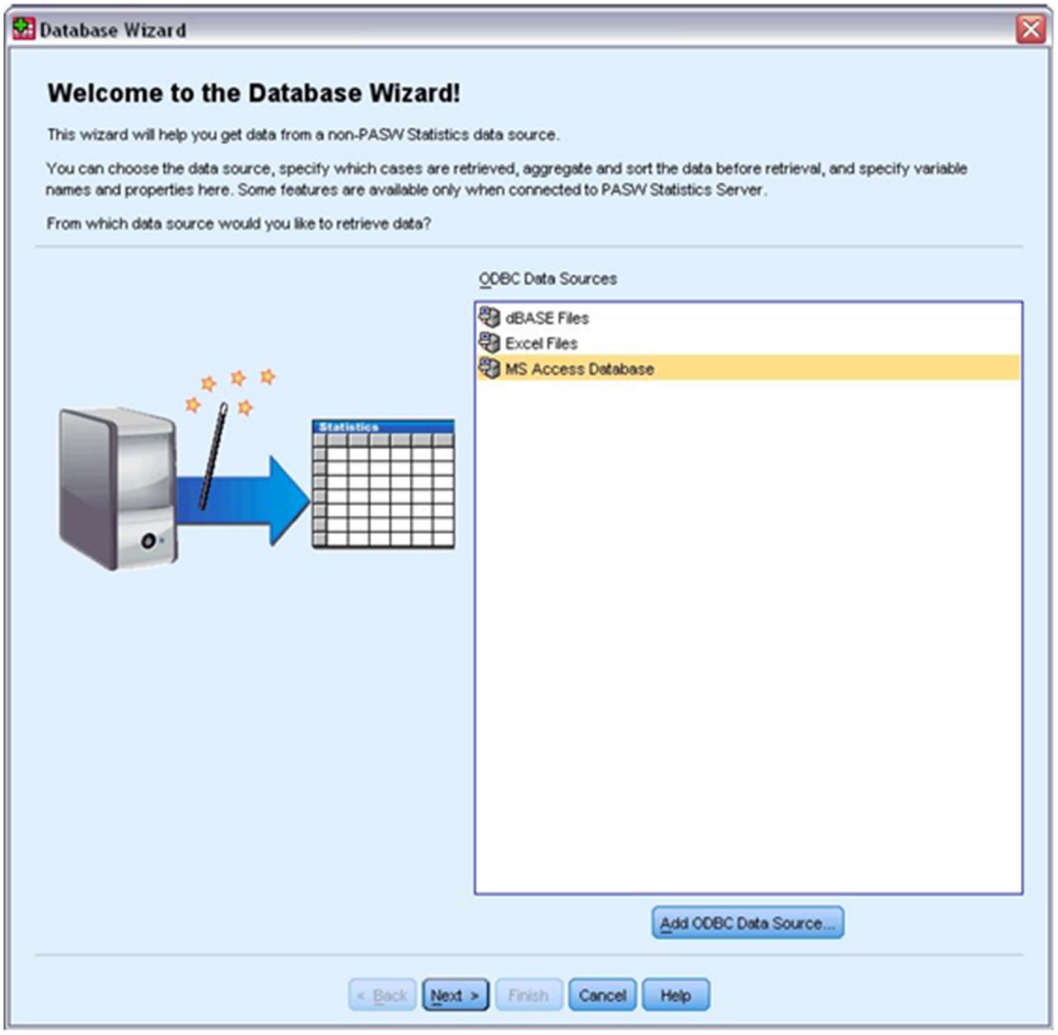

Database Wizard
Welcome to the Database Wizard!
This wizard will help you get data from a non-PASW Statistics data source.
You can choose the data source, specify which cases are retrieved, aggregate and sort the data before retrieval, and specify variable names and properties here. Some features are available only when connected to PASW Statistics Server.
From which data source would you like to retrieve data?
ODBC Data Sources
dBASE Files
Excel Files
MS Access Database
Statistics
Add ODBC Data Source...
< Back
Next >
Finish
Cancel
Help

Figure 13. Database Wizard Welcome dialog box

၂။ Data sources စာရင်းထဲကနေ MS Access Database ကိုရွေးပြီး Next ကိုနှိပ်ပါ။

မှတ်ချက် - သင့်ရဲ့ installation ပေါ်မူတည်ပြီး wizard ရဲ့ ဘယ်ဘက်အခြမ်းမှာ OLEDB data sources စာရင်းကိုလည်း မြင်ရနိုင်ပါတယ် (Windows operating systems တွေမှာသာ)။ ဒါပေမယ့် ဒီဥပမာမှာတော့ ညာဘက်အခြမ်းမှာ ပြထားတဲ့ ODBC data sources စာရင်းကိုပဲ သုံးသွားမှာပါ။

၃။ သင်ဖွင့်ချင်တဲ့ Access database file ဆီကိုသွားဖို့ Browse ကို နှိပ်ပါ။

၄။ demo.mdb ဖိုင်ကို ဖွင့်ပါ (ပိုသိချင်ရင် Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ ကြည့်နိုင်ပါတယ်)။

၅။ login dialog box မှာ OK ကို နှိပ်ပါ။ နောက်တစ်ဆင့်မှာတော့ သင် import လုပ်ချင်တဲ့ tables တွေနဲ့ variables တွေကို သတ်မှတ်ပေးနိုင်ပါတယ်။

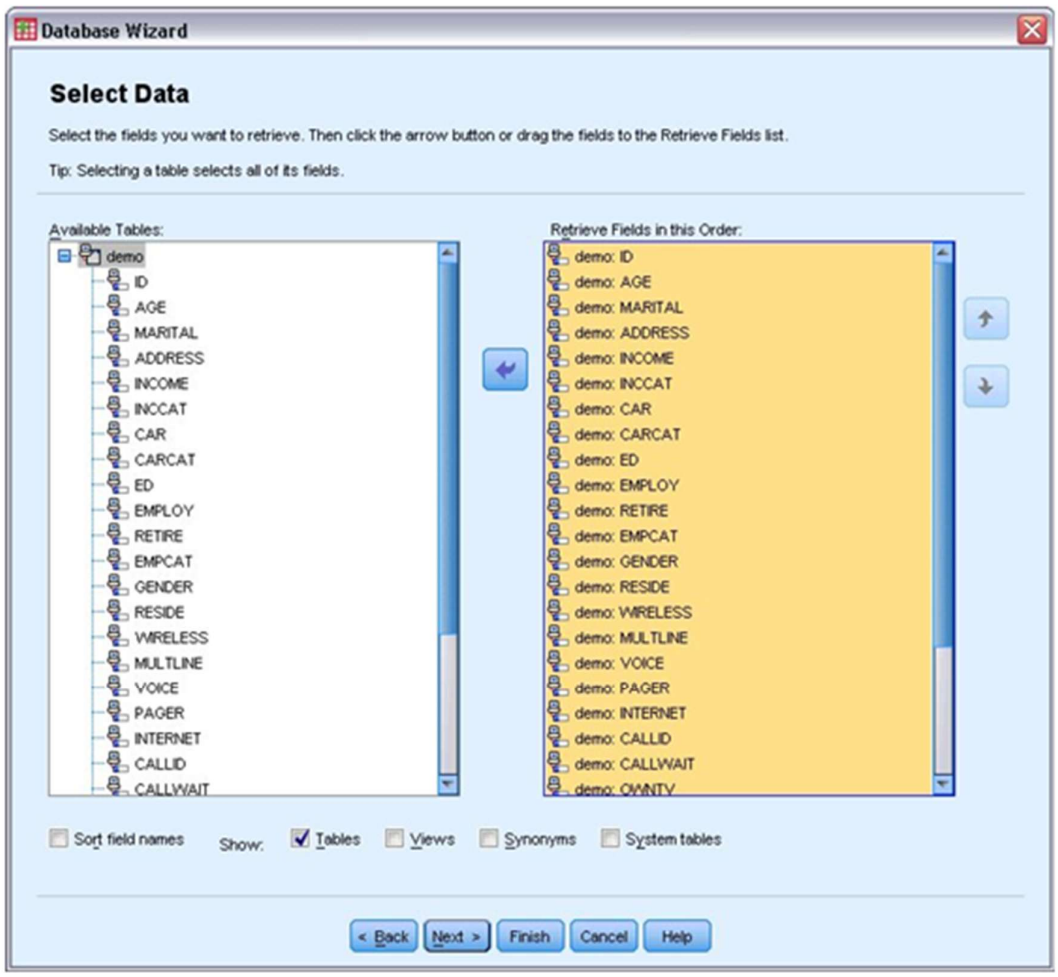


Figure 14. Select Data step

၆။ demo table တစ်ခုလုံးကို Retrieve Fields In This Order list ဆီကို ဆွဲထည့် (drag and drop) လိုက်ပါ။

၇။ Next ကို နှိပ်ပါ။

နောက်တစ်ဆင့်မှာ သင် ဘယ် records (cases) တွေကို import လုပ်မယ်ဆိုတာ ရွေးချယ်နိုင်ပါတယ်။ Cases အားလုံးကို import မလုပ်ချင်ဘူးဆိုရင်၊ cases ရဲ့ subset တစ်ခု (ဥပမာ၊ အသက် ၃၀ ကျော် အမျိုးသားများ) ကိုပဲဖြစ်ဖြစ်၊ ဒါမှမဟုတ် data source ထဲကနေ random sample တစ်ခုကိုပဲဖြစ်ဖြစ် import လုပ်နိုင်ပါတယ်။ ကြီးမားတဲ့ data sources တွေအတွက်ဆိုရင်တော့ processing time ကို လျှော့ချနိုင်ဖို့ cases အရေအတွက်ကို ကိုယ်စားပြုနိုင်တဲ့ sample အသေးလေးတစ်ခုလောက်ပဲ ကန့်သတ်ထားသင့်ပါတယ်။

၈။ ဆက်သွားဖို့ Next ကို နှိပ်ပါ။

Field names တွေကို variable names တွေ ဖန်တီးဖို့အတွက် အသုံးပြုပါတယ်။ လိုအပ်ရင် names တွေကို မှန်ကန်တဲ့ variable names တွေအဖြစ် ပြောင်းလဲပေးပါတယ်။ မူလ field names တွေကိုတော့ variable labels အနေနဲ့ ထိန်းသိမ်းထားပေးပါတယ်။ Database ကို import မလုပ်ခင် variable names တွေကို ပြင်ဆင်ပြောင်းလဲလို့လည်း ရပါသေးတယ်။

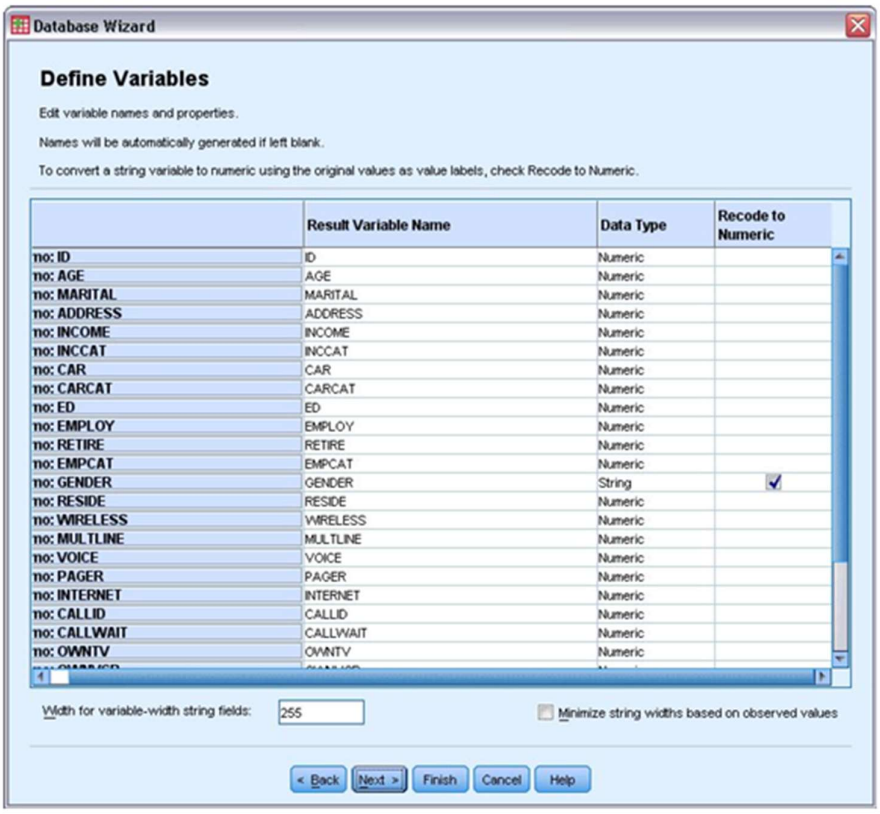

Database Wizard
Define Variables
Edit variable names and properties.
Names will be automatically generated if left blank.
To convert a string variable to numeric using the original values as value labels, check Recode to Numeric.
Result Variable Name
Data Type
Recode to Numeric
no: ID ID Numeric
no: AGE AGE Numeric
no: MARITAL MARITAL Numeric
no: ADDRESS ADDRESS Numeric
no: INCOME INCOME Numeric
no: INCCAT INCCAT Numeric
no: CAR CAR Numeric
no: CARCAT CARCAT Numeric
no: ED ED Numeric
no: EMPLOY EMPLOY Numeric
no: RETIRE RETIRE Numeric
no: EMPCAT EMPCAT Numeric
no: GENDER GENDER String
no: RESIDE RESIDE Numeric
no: WIRELESS WIRELESS Numeric
no: MULTLINE MULTLINE Numeric
no: VOICE VOICE Numeric
no: PAGER PAGER Numeric
no: INTERNET INTERNET Numeric
no: CALLID CALLID Numeric
no: CALLWAIT CALLWAIT Numeric
no: OWNTV OWNTV Numeric
Width for variable-width string fields:
255
Minimize string widths based on observed values
< Back
Next >
Finish
Cancel
Help

Figure 15. Define Variables step

၉။ Gender field ထဲက Recode to Numeric cell ကို ကလစ်နှိပ်ပါ။ ဒီ option က string variables တွေကို integer variables တွေအဖြစ် ပြောင်းလဲပေးပြီး မူလ value ကိုတော့ variable အသစ်အတွက် value label အနေနဲ့ ထိန်းသိမ်းထားပေးပါတယ်။

၁၀။ ဆက်သွားဖို့ Next ကို နှိပ်ပါ။

Database Wizard မှာ သင်ရွေးချယ်ခဲ့တဲ့အရာတွေကနေ ဖန်တီးလိုက်တဲ့ SQL statement က Results step မှာ ပေါ်လာပါလိမ့်မယ်။ ဒီ statement ကို အခုချက်ချင်း execute လုပ်လို့ရသလို၊ နောက်မှသုံးဖို့အတွက် ဖိုင်တစ်ခုအနေနဲ့ သိမ်းထားလို့လည်း ရပါတယ်။

၁၁။ Data ကို import လုပ်ဖို့ Finish ကို နှိပ်ပါ။

သင် import လုပ်ဖို့ ရွေးချယ်ခဲ့တဲ့ Access database ထဲက data အားလုံးကို အခု Data Editor ထဲမှာ အသုံးပြုနိုင်ပြီ ဖြစ်ပါတယ်။

Reading Data from a Text File (Text File မှ Data များကို ဖတ်ယူခြင်း)

Text files တွေကလည်း အများသုံး data source တစ်ခုဖြစ်ပါတယ်။ Spreadsheet programs နဲ့ databases တော်တော်များများက သူတို့ရဲ့ အချက်အလက်တွေကို text file formats အမျိုးမျိုးနဲ့ သိမ်းဆည်းပေးနိုင်ပါတယ်။ Comma- ဒါမှမဟုတ် tab-delimited files တွေဆိုတာက variable တစ်ခုစီကို ခွဲခြားပြသဖို့အတွက် commas တွေ ဒါမှမဟုတ် tabs တွေကို အသုံးပြုထားတဲ့ data rows တွေကို ဆိုလိုတာပါ။ ဒီဥပမာမှာတော့ data က tab-delimited ဖြစ်ပါတယ်။

၁။ menus တွေထဲကနေ File > Import Data > Text Data ကို ရွေးချယ်ပါ။

၂။ Samples\English folder ထဲကိုသွားပြီး demo.txt ကို ရွေးပါ။

Text Import Wizard က ရွေးချယ်ထားတဲ့ text file ကို ဘယ်လို interpret လုပ်မလဲ ဆိုတာ သတ်မှတ်ပေးတဲ့ လုပ်ငန်းစဉ်တစ်လျှောက် လမ်းညွှန်ပေးသွားမှာပါ။

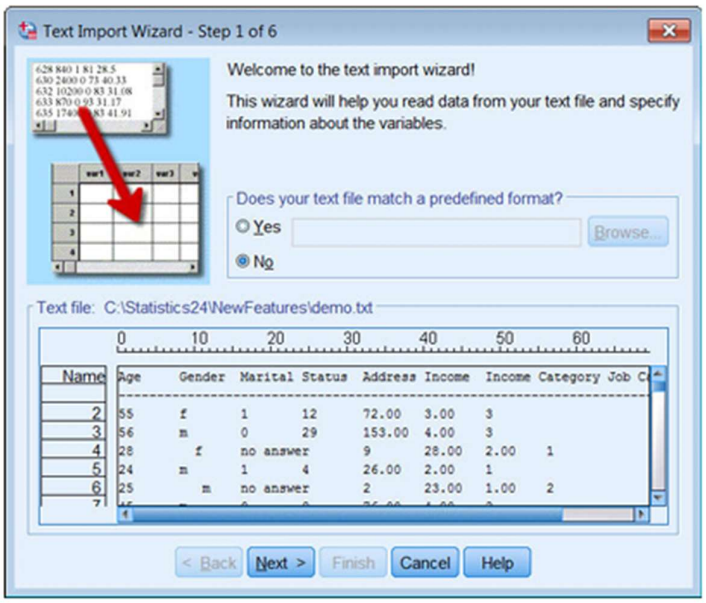


Figure 16. Text Import Wizard: Step 1 of 6

၃။ Step 1 မှာ သင်က predefined format (ကြိုတင်သတ်မှတ်ထားတဲ့ format) ကိုရွေးချယ်တာဖြစ်စေ၊ wizard ထဲမှာ format အသစ်တစ်ခု ဖန်တီးတာဖြစ်စေ လုပ်နိုင်ပါတယ်။ No ကို ရွေးချယ်ပါ။

၄။ ဆက်သွားဖို့ Next ကို နှိပ်ပါ။

စောစောက ပြောခဲ့သလိုပဲ၊ ဒီဖိုင်က tab-delimited formatting ကို သုံးထားတာပါ။ နောက်ပြီး variable names တွေကို ဒီဖိုင်ရဲ့ အပေါ်ဆုံးလိုင်းမှာ သတ်မှတ်ပေးထားပါတယ်။

၅။ wizard ရဲ့ step 2 မှာ၊ ဖိုင်က delimited formatting structure ကို သုံးထားတယ်ဆိုတာ ညွှန်ပြဖို့ Delimited ကို ရွေးချယ်ပါ။

၆။ ဖိုင်ရဲ့အပေါ်ဆုံးမှာ variable names တွေ ပါဝင်တယ်ဆိုတာကို ညွှန်ပြဖို့ Yes ကို ရွေးချယ်ပါ။

၇။ ဆက်သွားဖို့ Next ကို နှိပ်ပါ။

၈။ Step 3 မှာ၊ ပထမဆုံး data case စတင်တဲ့ line number အတွက် 2 လို့ ရိုက်ထည့်ပါ (ဘာလို့လဲဆိုတော့ variable names တွေက ပထမဆုံး line မှာ ရှိနေလို့ပါ)။

၉။ ဒီအဆင့်ရဲ့ ကျန်တဲ့နေရာတွေမှာ default values အတိုင်းပဲ ထားလိုက်ပြီး ဆက်သွားဖို့ Next ကို နှိပ်ပါ။ Step 4 မှာရှိတဲ့ Data preview က ဖိုင်ကို မှန်မှန်ကန်ကန် ဖတ်နိုင်ခြင်းရှိမရှိ အမြန်ဆုံး သေချာစေဖို့ လုပ်ဆောင်ပေးပါတယ်။

၁၀။ Tab ကို ရွေးချယ်ပြီး တခြား delimiters options တွေကို အမှန်ခြစ် ဖြုတ်လိုက်ပါ။ ဖိုင်ထဲမှာ spaces တွေ ပါနေတဲ့အတွက် Space ကို ပုံမှန်အားဖြင့် (by default) ရွေးချယ်ထားပါတယ်။ ဒီဖိုင်အတွက်ဆိုရင် spaces တွေက data values ရဲ့ အစိတ်အပိုင်းတစ်ခုဖြစ်ပြီး delimiters တွေ မဟုတ်ပါဘူး။ ဖိုင်ကို မှန်မှန်ကန်ကန် ဖတ်နိုင်ဖို့အတွက် Space မှာ အမှန်ခြစ် ဖြုတ်ပေးဖို့ လိုအပ်ပါတယ်။

၁၁။ Remove leading spaces for string values ကို ရွေးချယ်ပါ။ String values တွေရဲ့ အစမှာရှိနေတဲ့ spaces တွေက expressions တွေထဲမှာ string values တွေကို evaluate လုပ်တဲ့အခါ အကျိုးသက်ရောက်မှု ရှိနိုင်ပါတယ်။ ဒီဖိုင်မှာဆိုရင် Gender အတွက် တချို့ values တွေရဲ့ရှေ့မှာ value ရဲ့ အစိတ်အပိုင်းမဟုတ်တဲ့ spaces တွေ ပါနေပါတယ်။ အဲဒီ spaces တွေကိုသာ မဖယ်ရှားဘူးဆိုရင် " f" ဆိုတဲ့ value ဟာ "f" နဲ့ မတူညီတဲ့ value တစ်ခုအနေနဲ့ သတ်မှတ်ခံရမှာပါ။

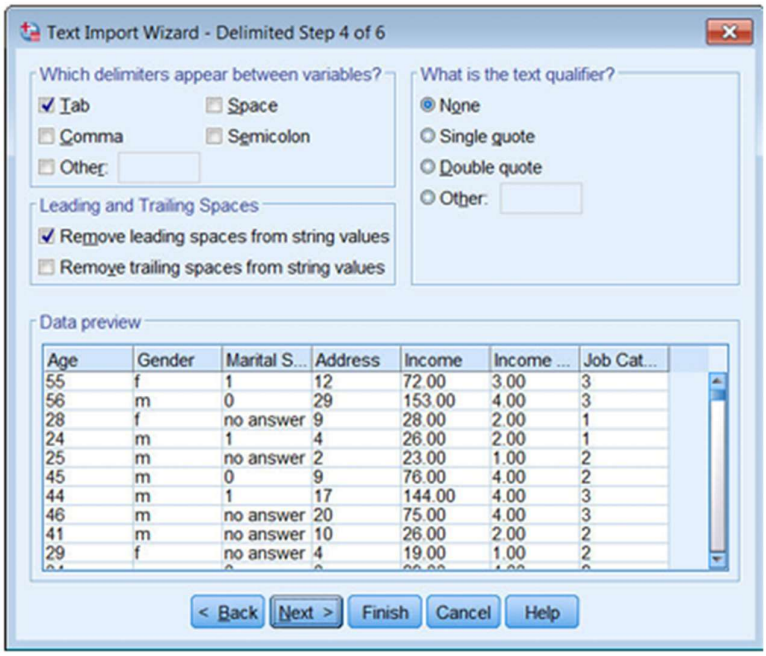


Figure 17. Text Import Wizard: Step 4 of 6

၁၂။ ဆက်သွားဖို့ Next ကို နှိပ်ပါ။

Variable names တွေကို naming rules တွေနဲ့ ကိုက်ညီအောင် ပြောင်းလဲထားတဲ့အတွက်၊ မလိုလားအပ်တဲ့ names တွေကို ပြင်ဆင်နိုင်ဖို့ step 5 မှာ အခွင့်အရေးပေးထားပါတယ်။

ဒီနေရာမှာ Data types တွေကိုလည်း သတ်မှတ်ပေးနိုင်ပါတယ်။ ဥပမာအနေနဲ့ Income ကို dollar currency format အဖြစ် ပြောင်းလဲနိုင်ပါတယ်။

Data type တစ်ခုကို ပြောင်းလဲဖို့အတွက်-

၁၃။ Data preview ထဲမှာ Income ကို ရွေးချယ်ပါ။

၁၄။ Data format drop-down list ထဲကနေ Dollar ကို ရွေးချယ်ပါ။

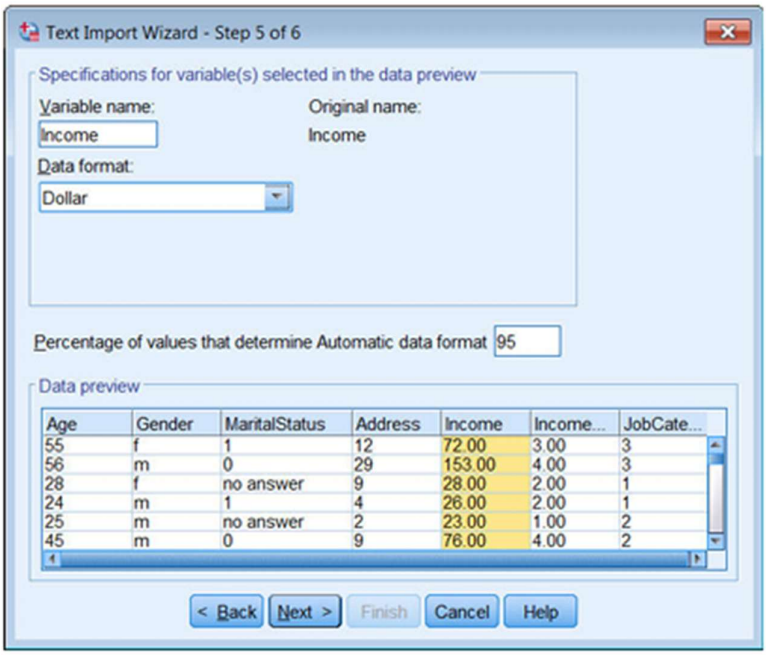

Text Import Wizard - Step 5 of 6
Specifications for variable(s) selected in the data preview
Variable name:
Income
Original name:
Income
Data format:
Dollar
Percentage of values that determine Automatic data format 95
Data preview
Age Gender MaritalStatus Address Income Income... JobCate...
55 f 1 12 72.00 3.00 3
56 m 0 29 153.00 4.00 3
28 f no answer 9 28.00 2.00 1
24 m 1 4 26.00 2.00 1
25 m no answer 2 23.00 1.00 2
45 m 0 9 76.00 4.00 2
< Back
Next >
Finish
Cancel
Help

Figure 18. Change the data type

MaritalStatus ဆိုတဲ့ variable မှာ string ရော numeric values တွေပါ နှစ်မျိုးလုံး ပါဝင်နေပါတယ်။ Values တွေရဲ့ ၅ ရာခိုင်နှုန်းအောက်ကသာ strings တွေ ဖြစ်ပါတယ်။ Automatic data format ကို သတ်မှတ်ပေးတဲ့ Percentage of values အတွက် default setting ဖြစ်တဲ့ 95% နဲ့ဆိုရင်၊ ဒီ variable ကို numeric အနေနဲ့ သတ်မှတ်ပြီး string values တွေကိုတော့ system-missing အဖြစ် သတ်မှတ်သွားမှာပါ။ တကယ်လို့ ဘယ် data format ကမှ အဲဒီ percentage value နဲ့ မကိုက်ညီဘူးဆိုရင်တော့ variable ကို string variable အနေနဲ့ သတ်မှတ်မှာပါ။ တကယ်လို့ သင်က setting ကို 100 လို့ ပြောင်းလိုက်မယ်ဆိုရင် values အားလုံးကို မူလအတိုင်း ထိန်းသိမ်းထားပေးမှာဖြစ်ပေမယ့် numeric values အားလုံးကိုတော့ strings တွေအဖြစ် သတ်မှတ်သွားမှာ ဖြစ်ပါတယ်။

၁၅။ ဆက်သွားဖို့ Next ကို နှိပ်ပါ။

၁၆။ နောက်ဆုံးအဆင့်မှာတော့ default selections အတိုင်းပဲ ထားလိုက်ပြီး data ကို import လုပ်ဖို့ Finish ကို နှိပ်ပါ။

Chapter 3. Using the Data Editor

Data Editor ဟာ active ဖြစ်နေတဲ့ data file ရဲ့ အကြောင်းအရာတွေကို ပြသပေးပါတယ်။ Data Editor ထဲက သတင်းအချက်အလက်တွေကို variables တွေနဲ့ cases တွေအနေနဲ့ ဖွဲ့စည်းထားတာပါ။

- Data View မှာဆိုရင် columns တွေက variables တွေကို ကိုယ်စားပြုပြီး၊ rows တွေက cases (observations) တွေကို ကိုယ်စားပြုပါတယ်။
- Variable View မှာတော့ row တစ်ခုစီဟာ variable တစ်ခုဖြစ်ပြီး၊ column တစ်ခုစီဟာ အဲဒီ variable နဲ့ သက်ဆိုင်တဲ့ attribute တစ်ခုဖြစ်ပါတယ်။

Variables တွေကို သင် စုစည်းထားတဲ့ မတူညီတဲ့ data အမျိုးအစားတွေကို ကိုယ်စားပြုဖို့ အသုံးပြုပါတယ်။ အလွယ်ဆုံး ဥပမာပေးရရင် survey တစ်ခုလိုပေါ့။ Survey မှာပါတဲ့ မေးခွန်းတစ်ခုစီရဲ့ အဖြေဟာ variable တစ်ခုနဲ့ ညီမျှပါတယ်။ Variables တွေမှာ numbers, strings, currency နဲ့ dates စသဖြင့် အမျိုးအစားစုံလင်စွာ ရှိပါတယ်။

Entering Numeric Data

Data တွေကို Data Editor ထဲ တိုက်ရိုက် ရိုက်ထည့်လို့ရပါတယ်၊ ဒီနည်းလမ်းက သေးငယ်တဲ့ data files တွေအတွက် ဒါမှမဟုတ် ကြီးမားတဲ့ data files တွေမှာ အသေးစား ပြင်ဆင်မှုလေးတွေ လုပ်တဲ့အခါ အသုံးဝင်ပါတယ်။

၁။ Data Editor window ရဲ့ အောက်ခြေမှာရှိတဲ့ Variable View tab ကို နှိပ်ပါ။ သင် အသုံးပြုမယ့် variables တွေကို အရင်ဆုံး သတ်မှတ် (define) ပေးဖို့ လိုပါတယ်။ ဒီနေရာမှာတော့ age, marital status နဲ့ income ဆိုတဲ့ variables သုံးခုပဲ လိုအပ်ပါတယ်။

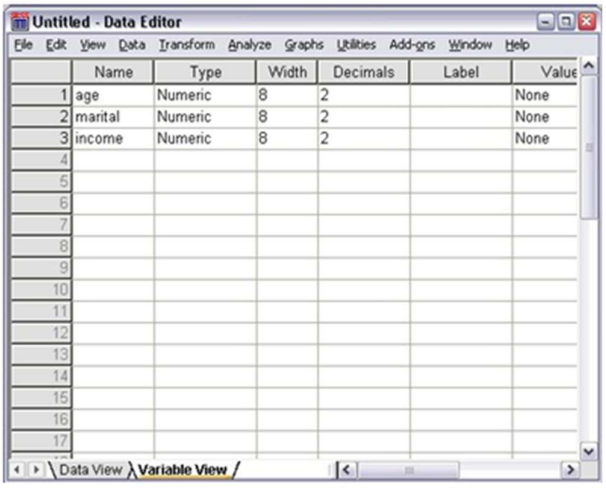

Untitled - Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
Name Type Width Decimals Label Value
1 age Numeric 8 2 None
2 marital Numeric 8 2 None
3 income Numeric 8 2 None
Data View Variable View

Figure 19. Variable names in Variable View

၂။ ပထမဆုံး row ရဲ့ ပထမဆုံး column မှာ age လို့ ရိုက်ထည့်ပါ။

၃။ ဒုတိယ row မှာ marital လို့ ရိုက်ထည့်ပါ။

၄။ တတိယ row မှာ income လို့ ရိုက်ထည့်ပါ။

Variable အသစ်တွေကို ပုံမှန်အားဖြင့် Numeric data type အနေနဲ့ အလိုအလျောက် သတ်မှတ်ပေးပါတယ်။

တကယ်လို့ သင်က variable names တွေကို မရိုက်ထည့်ဘူးဆိုရင်လည်း unique names တွေကို အလိုအလျောက် ဖန်တီးပေးမှာပါ။ ဒါပေမယ့် အဲဒီ names တွေက အဓိပ္ပာယ်ပြည့်စုံမှုမရှိတဲ့အတွက် ကြီးမားတဲ့ data files တွေမှာ သုံးဖို့ အကြံမပြုပါဘူး။

၅။ Data တွေကို ဆက်ပြီး ရိုက်ထည့်ဖို့ Data View tab ကို နှိပ်ပါ။

Variable View မှာ သင်ရိုက်ထည့်ခဲ့တဲ့ names တွေက အခုဆိုရင် Data View ရဲ့ ပထမဆုံး columns သုံးခုအတွက် headings တွေ ဖြစ်နေပါပြီ။ ပထမဆုံး row ရဲ့ ပထမဆုံး column ကနေစပြီး data တွေကို စတင်ရိုက်ထည့်ပါ။

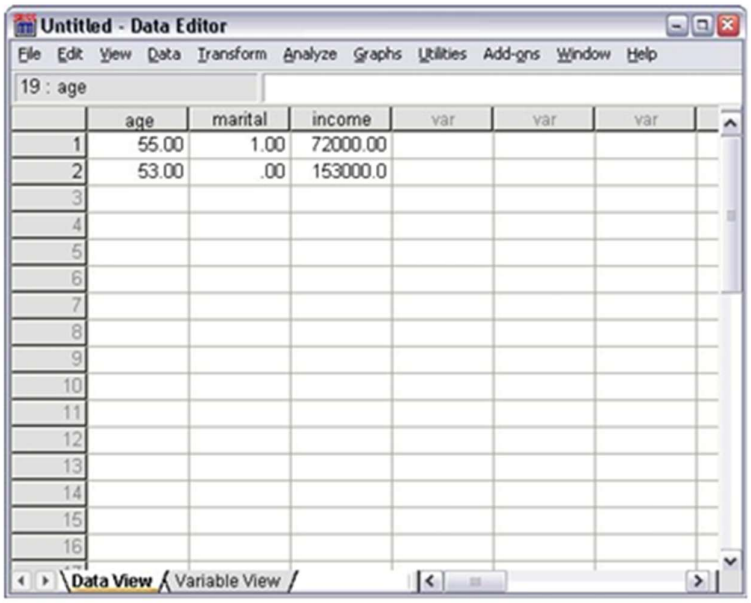

Untitled - Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
19 : age
age marital income var var var
1 55.00 1.00 72000.00
2 53.00 .00 153000.0
Data View Variable View

Figure 20. Values entered in Data View

၆။ age column မှာ 55 လို့ ရိုက်ထည့်ပါ။

၇။ marital column မှာ 1 လို့ ရိုက်ထည့်ပါ။

၈။ income column မှာ 72000 လို့ ရိုက်ထည့်ပါ။

၉။ နောက်ထပ် subject တစ်ခုရဲ့ data ကို ထည့်ဖို့အတွက် cursor ကို ဒုတိယ row ရဲ့ ပထမဆုံး column ဆီ ရွှေ့ပါ။

၁၀။ age column မှာ 53 လို့ ရိုက်ထည့်ပါ။

၁၁။ marital column မှာ 0 လို့ ရိုက်ထည့်ပါ။

၁၂။ income column မှာ 153000 လို့ ရိုက်ထည့်ပါ။

လောလောဆယ်မှာတော့ age နဲ့ marital columns တွေဟာ integers တွေဖြစ်ဖို့ ရည်ရွယ်ထားပေမယ့် decimal points တွေ ပြသနေပါတယ်။ အဲဒီ variables တွေမှာ decimal points တွေကို ဖျောက်ချင်တယ်ဆိုရင်-

၁၃။ Data Editor window အောက်ခြေက Variable View tab ကို နှိပ်ပါ။

၁၄။ age row ရဲ့ Decimals column မှာ ဒသမကို ဖျောက်ဖို့ 0 လို့ ရိုက်ထည့်ပါ။

၁၅။ marital row ရဲ့ Decimals column မှာ ဒသမကို ဖျောက်ဖို့ 0 လို့ ရိုက်ထည့်ပါ။

Entering String Data

ကိန်းဂဏန်းမဟုတ်တဲ့ စာသားတွေကိုလည်း Data Editor ထဲမှာ ထည့်သွင်းနိုင်ပါတယ်။

၁. Data Editor window အောက်ခြေမှာရှိတဲ့ Variable View tab ကိုနှိပ်ပါ။

၂. row အလွတ်ရဲ့ ပထမဆုံး cell ထဲမှာ variable name အဖြစ် sex ကို ရိုက်ပါ။

၃. Type ကိုနှိပ်ပါ။

၄. Variable Type dialog box ကိုဖွင့်ဖို့ type ရဲ့ ညာဘက်က button ကိုနှိပ်ပါ။

၅. variable type ကို String ကို ရွေးပေးပါ။

၆. OK ကိုနှိပ်ပါ။ Data Editor ဆီကို ပြန်သွားလို့ရပါပြီ။

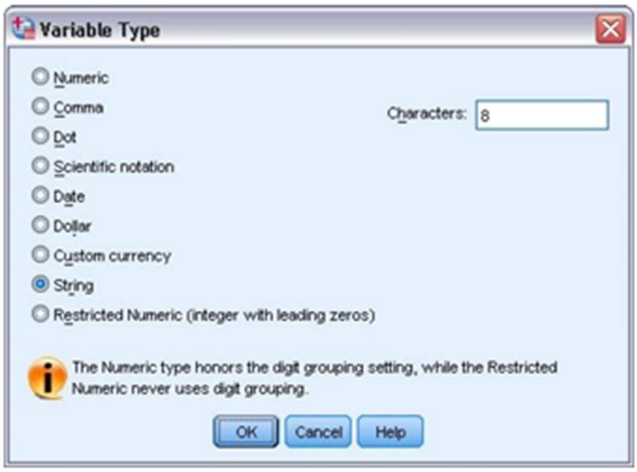


Figure 21. Variable Type dialog box

Defining Data

Data types တွေကို သတ်မှတ်တဲ့အပြင်၊ variable names တွေနဲ့ data values တွေအတွက် အသေးစိတ်ကျတဲ့ variable labels တွေနဲ့ value labels တွေကိုလည်း သတ်မှတ်ပေးနိုင်ပါတယ်။ ဒီ descriptive labels တွေကို statistical reports တွေနဲ့ charts တွေမှာ အသုံးပြုကြပါတယ်။

Adding Variable Labels

Labels ဆိုတာ variables တွေရဲ့ အကြောင်းအရာကို ဖော်ပြဖို့ ရည်ရွယ်ပါတယ်။ ဒီ descriptions တွေက များသောအားဖြင့် variable names တွေထက် ပိုရှည်တတ်ပါတယ်။ Labels တွေကို 255 bytes အထိ ပေးလို့ရပါတယ်။ ဒီ labels တွေကို output ထုတ်တဲ့အခါ မတူညီတဲ့ variables တွေကို ခွဲခြားသိမြင်ဖို့ အသုံးပြုပါတယ်။

၁။ Data Editor window အောက်ခြေက Variable View tab ကို နှိပ်ပါ။

၂။ age row ရဲ့ Label column မှာ Respondent's Age လို့ ရိုက်ထည့်ပါ။

၃။ marital row ရဲ့ Label column မှာ Marital Status လို့ ရိုက်ထည့်ပါ။

၄။ income row ရဲ့ Label column မှာ Household Income လို့ ရိုက်ထည့်ပါ။

၅။ sex row ရဲ့ Label column မှာ Gender လို့ ရိုက်ထည့်ပါ။

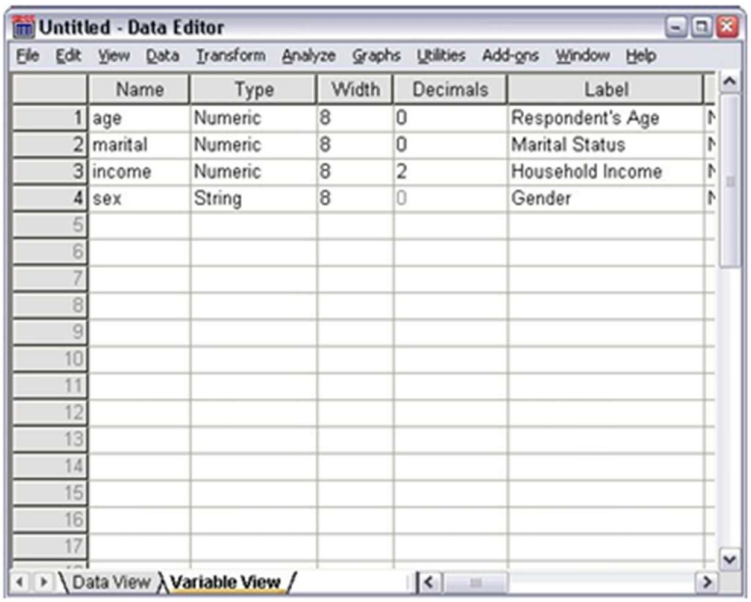

Untitled - Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
Name Type Width Decimals Label
1 age Numeric 8 0 Respondent's Age
2 marital Numeric 8 0 Marital Status
3 income Numeric 8 2 Household Income
4 sex String 8 0 Gender
Data View Variable View

Figure 22. Variable labels entered in Variable View

Changing Variable Type and Format

Type column က variable တစ်ခုစီရဲ့ လက်ရှိ data type ကို ပြသပေးပါတယ်။ အသုံးအများဆုံး data types တွေကတော့ numeric နဲ့ string တွေဖြစ်ပေမယ့် တခြား formats အများကြီးကိုလည်း support လုပ်ပေးပါသေးတယ်။ လက်ရှိ data file မှာဆိုရင် income variable ကို numeric type အနေနဲ့ သတ်မှတ်ထားပါတယ်။

၁။ income row အတွက် Type cell ကို နှိပ်ပါ။ ပြီးရင် Variable Type dialog box ကို ဖွင့်ဖို့ cell ရဲ့ ညာဘက်ဘေးက button လေးကို နှိပ်ပါ။

၂။ Dollar ကို ရွေးချယ်ပါ။

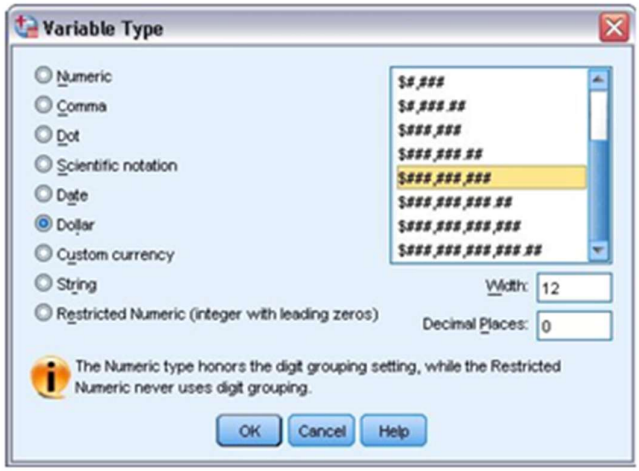


Figure 23. Variable Type dialog box

လက်ရှိ ရွေးချယ်ထားတဲ့ data type အတွက် formatting options တွေ ပေါ်လာပါလိမ့်မယ်။

၃။ ဒီဥပမာမှာပါတဲ့ currency ရဲ့ format အတွက် $###,###,### ကို ရွေးချယ်ပါ။

၄။ ပြင်ဆင်မှုတွေကို သိမ်းဆည်းဖို့ OK ကို နှိပ်ပါ။

Adding Value Labels

Value labels ဆိုတာက သင်ရဲ့ variable values တွေကို string label တစ်ခုနဲ့ ချိတ်ဆက် (map) ပေးတဲ့ နည်းလမ်းတစ်ခု ဖြစ်ပါတယ်။ ဒီဥပမာမှာ marital variable အတွက် လက်ခံနိုင်တဲ့ values နှစ်ခု ရှိပါတယ်။ Value 0 ဆိုတာက subject ဟာ single ဖြစ်တယ်လို့ အဓိပ္ပာယ်ရပြီး၊ Value 1 ဆိုတာကတော့ married ဖြစ်တယ်လို့ ဆိုလိုပါတယ်။

၁။ marital row အတွက် Values cell ကို နှိပ်ပါ၊ ပြီးရင် Value Labels dialog box ကို ဖွင့်ဖို့ cell ရဲ့ ညာဘက်ဘေးက button လေးကို နှိပ်ပါ။

Value ဆိုတာက အမှန်တကယ် numeric value ပါ။

Value label ကတော့ သတ်မှတ်ထားတဲ့ numeric value အတွက် အသုံးပြုမယ့် string label ဖြစ်ပါတယ်။

၂။ Value field မှာ 0 လို့ ရိုက်ထည့်ပါ။

၃။ Label field မှာ Single လို့ ရိုက်ထည့်ပါ။

၄။ ဒီ label ကို list ထဲ ထည့်ဖို့ Add ကို နှိပ်ပါ။

Figure 24. Value Labels dialog box

၅။ Value field မှာ 1 လို့ ရိုက်ထည့်ပြီး၊ Label field မှာ Married လို့ ရိုက်ထည့်ပါ။

၆။ Add ကို နှိပ်ပါ၊ ပြီးရင် ပြင်ဆင်မှုတွေကို သိမ်းဆည်းပြီး Data Editor ဆီ ပြန်သွားဖို့ OK ကို နှိပ်ပါ။

ဒီ labels တွေကို Data View မှာလည်း ပြသနိုင်တဲ့အတွက် သင့်ရဲ့ data တွေကို ပိုပြီး ဖတ်ရလွယ်ကူစေပါတယ်။

၇။ Data Editor window အောက်ခြေက Data View tab ကို နှိပ်ပါ။

၈။ menus တွေထဲကနေ View > Value Labels ကို ရွေးချယ်ပါ။

အခုဆိုရင် Data Editor ထဲမှာ values တွေ ရိုက်ထည့်တဲ့အခါ labels တွေကို list တစ်ခုအနေနဲ့ ပြသပေးနေပြီ ဖြစ်ပါတယ်။

ဒီလိုလုပ်ထားခြင်းအားဖြင့် မှန်ကန်တဲ့ အဖြေကို အကြံပြုပေးနိုင်ပြီး၊ ပိုပြီး ပြည့်စုံတဲ့ အဖြေတစ်ခုကို ပေးနိုင်တဲ့ အကျိုးကျေးဇူး ရှိပါတယ်။

တကယ်လို့ Value Labels menu item က ပွင့်ပြီးသား ဖြစ်နေမယ်ဆိုရင် (ဘေးမှာ အမှန်ခြစ်လေး ရှိနေရင်)၊ Value Labels ကို ထပ်ရွေးလိုက်တာဟာ value labels တွေ ပြသနေတာကို ပိတ်လိုက်မှာ ဖြစ်ပါတယ်။

Handling Missing Data

Missing ဒါမှမဟုတ် invalid ဖြစ်နေတဲ့ data တွေဟာ လျစ်လျူရှုထားဖို့ မဖြစ်နိုင်လောက်အောင်ကို အတွေ့ရများပါတယ်။ Survey ဖြေဆိုသူတွေက တချို့မေးခွန်းတွေကို ဖြေဖို့ ငြင်းဆန်တာ၊ အဖြေကို မသိတာ၊ ဒါမှမဟုတ် မမျှော်လင့်ထားတဲ့ format မျိုးနဲ့ ဖြေဆိုတာတွေ ရှိနိုင်ပါတယ်။ ဒီ data တွေကိုသာ filter မလုပ်ဘူး၊ ခွဲခြားမသတ်မှတ်ဘူးဆိုရင်၊ သင့်ရဲ့ analysis က တိကျတဲ့ results တွေကို ပေးနိုင်မှာ မဟုတ်ပါဘူး။

Numeric data တွေအတွက်ဆိုရင်၊ ဘာမှမရေးထားတဲ့ ကွက်လပ်တွေ ဒါမှမဟုတ် မှားယွင်းတဲ့ အချက်အလက်တွေ ပါနေတဲ့ fields တွေကို system-missing အဖြစ် ပြောင်းလဲပေးလိုက်ပြီး၊ ၎င်းကို အစက်ကလေး (period) တစ်စက်တည်းနဲ့ သိနိုင်ပါတယ်။

Value တစ်ခု ဘာကြောင့် missing ဖြစ်နေရသလဲဆိုတဲ့ အကြောင်းရင်းက သင့်ရဲ့ analysis အတွက် အရေးကြီးနိုင်ပါတယ်။ ဥပမာအားဖြင့်၊ မေးခွန်းကို ဖြေဖို့ ငြင်းဆန်သွားတဲ့ ဖြေဆိုသူတွေနဲ့ သူတို့နဲ့ မသက်ဆိုင်လို့ မဖြေခဲ့တဲ့ ဖြေဆိုသူတွေကြား ခွဲခြားသိမြင်ဖို့က အသုံးဝင်နိုင်ပါတယ်။

Missing Values for a Numeric Variable

၁။ Data Editor window အောက်ခြေက Variable View tab ကို နှိပ်ပါ။

၂။ age row မှာရှိတဲ့ Missing cell ကို နှိပ်ပါ၊ ပြီးရင် Missing Values dialog box ကို ဖွင့်ဖို့ cell ရဲ့ ညာဘက်ဘေးက button လေးကို နှိပ်ပါ။

ဒီ dialog box ထဲမှာ မတူညီတဲ့ missing values သုံးခုအထိကို သတ်မှတ်နိုင်သလို၊ values range တစ်ခုနဲ့အတူ နောက်ထပ် discrete value တစ်ခုကိုလည်း သတ်မှတ်ပေးနိုင်ပါတယ်။

Figure 25. Missing Values dialog box

၃။ Discrete missing values ကို ရွေးချယ်ပါ။

၄။ ပထမဆုံး text box မှာ 999 လို့ ရိုက်ထည့်ပြီး ကျန်တဲ့ text boxes နှစ်ခုကို အလွတ်ထားခဲ့ပါ။

၅။ ပြင်ဆင်မှုတွေကို သိမ်းဆည်းပြီး Data Editor ဆီ ပြန်သွားဖို့ OK ကို နှိပ်ပါ။

အခု missing data value ကို ထည့်သွင်းပြီးသွားပြီဖြစ်တဲ့အတွက်၊ အဲဒီ value အတွက် label တစ်ခု သတ်မှတ်ပေးလို့ ရပါပြီ။

၆။ age row ရဲ့ Values cell ကို နှိပ်ပါ၊ ပြီးရင် Value Labels dialog box ကို ဖွင့်ဖို့ cell ရဲ့ ညာဘက်ဘေးက button လေးကို နှိပ်ပါ။

၇။ Value field မှာ 999 လို့ ရိုက်ထည့်ပါ။

၈။ Label field မှာ No Response လို့ ရိုက်ထည့်ပါ။

၉။ ဒီ label ကို data file ထဲ ထည့်ဖို့ Add ကို နှိပ်ပါ။

၁၀။ ပြင်ဆင်မှုတွေကို သိမ်းဆည်းပြီး Data Editor ဆီ ပြန်သွားဖို့ OK ကို နှိပ်ပါ။

Missing Values for a String Variable

String variables တွေအတွက် missing values တွေကိုလည်း numeric variables တွေရဲ့ missing values တွေနဲ့ ဆင်တူစွာ ကိုင်တွယ်ဖြေရှင်းနိုင်ပါတယ်။ ဒါပေမယ့် numeric variables တွေနဲ့ မတူတာကတော့၊ string variables တွေက အလွတ်ဖြစ်နေတဲ့ fields တွေကို system-missing အနေနဲ့ သတ်မှတ်မထားပါဘူး။ ၎င်းတို့ကို empty string အနေနဲ့ပဲ မှတ်ယူပါတယ်။

၁။ Data Editor window အောက်ခြေက Variable View tab ကို နှိပ်ပါ။

၂။ sex row မှာရှိတဲ့ Missing cell ကို နှိပ်ပါ၊ ပြီးရင် Missing Values dialog box ကို ဖွင့်ဖို့ cell ရဲ့ ညာဘက်ဘေးက button လေးကို နှိပ်ပါ။

၃။ Discrete missing values ကို ရွေးချယ်ပါ။

၄။ ပထမဆုံး text box မှာ NR လို့ ရိုက်ထည့်ပါ။

String variables တွေအတွက် missing values တွေက case sensitive ဖြစ်ပါတယ်။ ဒါကြောင့် nr ဆိုတဲ့ value လေးကို missing value အဖြစ် သတ်မှတ်မှာ မဟုတ်ပါဘူး။

၅။ ပြင်ဆင်မှုတွေကို သိမ်းဆည်းပြီး Data Editor ဆီ ပြန်သွားဖို့ OK ကို နှိပ်ပါ။

အခုဆိုရင် missing value အတွက် label တစ်ခုကို ထပ်ထည့်နိုင်ပါပြီ။

၆။ sex row ရဲ့ Values cell ကို နှိပ်ပါ၊ ပြီးရင် Value Labels dialog box ကို ဖွင့်ဖို့ cell ရဲ့ ညာဘက်ဘေးက button လေးကို နှိပ်ပါ။

၇။ Value field မှာ NR လို့ ရိုက်ထည့်ပါ။

၈။ Label field မှာ No Response လို့ ရိုက်ထည့်ပါ။

၉။ ဒီ label ကို project ထဲ ထည့်ဖို့ Add ကို နှိပ်ပါ။

၁၀။ ပြင်ဆင်မှုတွေကို သိမ်းဆည်းပြီး Data Editor ဆီ ပြန်သွားဖို့ OK ကို နှိပ်ပါ။

Chapter 4. Examining Summary Statistics for Individual Variables

ဒီ chapter မှာ ရိုးရှင်းတဲ့ summary measures တွေအကြောင်းနဲ့ variable တစ်ခုရဲ့ level of measurement က အသုံးပြုသင့်တဲ့ statistics အမျိုးအစားတွေအပေါ် ဘယ်လိုလွှမ်းမိုးမှုရှိသလဲဆိုတာကို ဆွေးနွေးသွားမှာပါ။

ဒီဥပမာတွေအတွက် demo.sav ဆိုတဲ့ data file ကို အသုံးပြုသွားမှာပါ။ ပိုပြီးသိချင်တယ်ဆိုရင်တော့ Chapter 10 က "Sample Files" ခေါင်းစဉ်မှာ သွားရောက်လေ့လာနိုင်ပါတယ်။

Level of Measurement

Data အမျိုးအစားတွေအပေါ် မူတည်ပြီး သူတို့ရဲ့ level of measurement နဲ့ ကိုက်ညီမယ့် မတူညီတဲ့ summary measures တွေကို အသုံးပြုဖို့ လိုအပ်ပါတယ်-

- **Categorical**. ကွဲပြားတဲ့ values ဒါမှမဟုတ် categories အရေအတွက် ကန့်သတ်ချက်ရှိတဲ့ data မျိုးကို ဆိုလိုတာပါ (ဥပမာအားဖြင့် gender ဒါမှမဟုတ် marital status လိုမျိုးပေါ့)။ သူ့ကို qualitative data လို့လည်း ခေါ်ကြပါတယ်။ Categorical variables တွေဟာ string (alphanumeric) data တွေ ဖြစ်နိုင်သလို၊ categories တွေကို ကိုယ်စားပြုဖို့ numeric codes တွေသုံးထားတဲ့ numeric variables တွေလည်း ဖြစ်နိုင်ပါတယ် (ဥပမာ၊ 0 = Unmarried နဲ့ 1 = Married ဆိုပြီး သတ်မှတ်တာမျိုးပါ)။ Categorical data မှာ အခြေခံအားဖြင့် အမျိုးအစား နှစ်မျိုး ရှိပါတယ်-
    - **Nominal**. Categories တွေကြားမှာ မွေးရာပါ အစီအစဉ် (inherent order) သတ်မှတ်ထားခြင်းမရှိတဲ့ categorical data မျိုးပါ။ ဥပမာအားဖြင့် sales ဆိုတဲ့ job category ဟာ marketing ဒါမှမဟုတ် research ဆိုတဲ့ job category တွေထက် ပိုမြင့်တယ်၊ ပိုနိမ့်တယ်လို့ သတ်မှတ်လို့မရပါဘူး။
    - **Ordinal**. Categories တွေကြားမှာ အဓိပ္ပာယ်ရှိတဲ့ အစီအစဉ်တစ်ခု ရှိပေမယ့်၊ categories တစ်ခုနဲ့တစ်ခုကြားက အကွာအဝေး (distance) ကို တိုင်းတာလို့မရတဲ့ categorical data မျိုးဖြစ်ပါတယ်။ ဥပမာ၊ high, medium, low ဆိုတဲ့ values တွေမှာ အစီအစဉ်တစ်ခု ရှိနေပေမယ့် အဲဒီ values တွေကြားက "distance" ကိုတော့ တွက်ချက်လို့ မရပါဘူး။
- **Scale**. Interval ဒါမှမဟုတ် ratio scale ပေါ်မှာ တိုင်းတာထားတဲ့ data ဖြစ်ပြီး၊ data values တွေက values တွေရဲ့ အစီအစဉ် (order) ရော၊ values တွေကြားက အကွာအဝေး (distance) ကိုပါ ညွှန်ပြနိုင်ပါတယ်။ ဥပမာအနေနဲ့ လစာ $72,195 ဟာ $52,398 ထက်

ပိုများပြီး၊ အဲဒီ values နှစ်ခုကြားက distance က $19,797 ဖြစ်တယ်ဆိုတာ တွက်ထုတ်လို့ရပါတယ်။ သူ့ကို quantitative data ဒါမှမဟုတ် continuous data လို့လည်း ခေါ်ဆိုကြပါတယ်။

Summary Measures for Categorical Data

Categorical data တွေအတွက် အသုံးအများဆုံး summary measure ကတော့ category တစ်ခုစီမှာရှိတဲ့ cases အရေအတွက် (number) ဒါမှမဟုတ် ရာခိုင်နှုန်း (percentage) ပဲ ဖြစ်ပါတယ်။ Mode ဆိုတာကတော့ cases အရေအတွက် အများဆုံးရှိတဲ့ category ကို ခေါ်တာပါ။ Ordinal data တွေအတွက်ဆိုရင်တော့ categories အရေအတွက် အများကြီးရှိနေတဲ့အခါ median (cases တွေရဲ့ တစ်ဝက်က အထက်မှာရှိပြီး ကျန်တစ်ဝက်က အောက်မှာရှိနေတဲ့ value) ဟာလည်း အသုံးဝင်တဲ့ summary measure တစ်ခု ဖြစ်လာနိုင်ပါတယ်။

Frequencies procedure က variable တစ်ခုရဲ့ observed value တစ်ခုချင်းစီအတွက် cases အရေအတွက်ရော ရာခိုင်နှုန်းကိုပါ ပြသပေးတဲ့ frequency tables တွေကို ထုတ်ပေးပါတယ်။

၁။ menus တွေထဲကနေ Analyze > Descriptive Statistics > Frequencies... ကို ရွေးချယ်ပါ။ (မှတ်ချက် - ဒီ feature ကိုသုံးဖို့ Statistics Base option လိုအပ်ပါတယ်။)

၂။ Owns PDA [ownpda] နဲ့ Owns TV [owntv] ကို ရွေးချယ်ပြီး Variable(s) list ထဲကို ရွှေ့ထည့်ပါ။

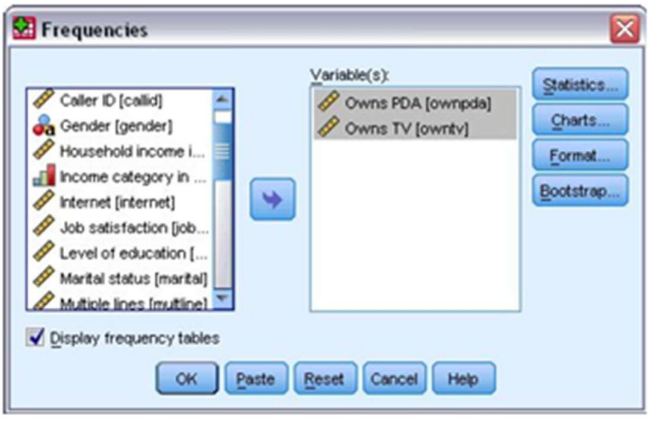


Figure 26. Categorical variables selected for analysis

၃။ procedure ကို run ဖို့အတွက် OK ကို နှိပ်ပါ။

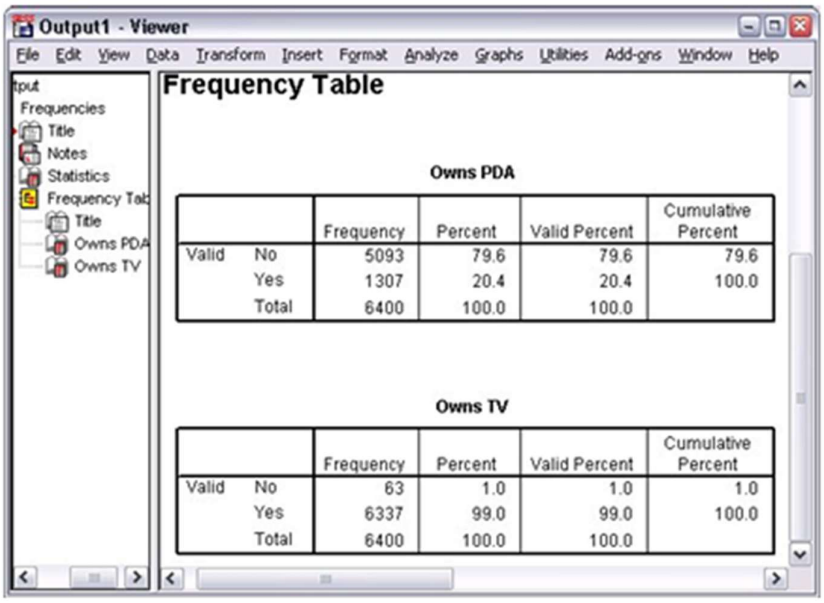


Figure 27. Frequency tables

Frequency tables တွေကို Viewer window ထဲမှာ ပြသပေးပါလိမ့်မယ်။ ဒီ frequency tables တွေအရ လူတွေရဲ့ 20.4% ကသာ PDAs တွေကို ပိုင်ဆိုင်ကြပေမယ့်၊ လူတိုင်းနီးပါး (99.0%) က TV ပိုင်ဆိုင်ကြတယ်ဆိုတာကို တွေ့ရပါတယ်။ ဒီအချက်တွေက သိပ်ပြီး စိတ်ဝင်စားစရာကောင်းတဲ့ တွေ့ရှိချက်တွေ မဟုတ်နိုင်ပေမယ့်၊ TV မပိုင်ဆိုင်တဲ့ လူစုလေးနဲ့ပတ်သက်ပြီး ပိုမိုလေ့လာရှာဖွေကြည့်တာကတော့ စိတ်ဝင်စားစရာ ကောင်းနိုင်ပါတယ်။

Charts for Categorical Data

Frequency table ထဲက အချက်အလက်တွေကို bar chart ဒါမှမဟုတ် pie chart တွေသုံးပြီး graphically ပြသနိုင်ပါတယ်။

၁။ Frequencies dialog box ကို ပြန်ဖွင့်ပါ (စောစောက variables နှစ်ခုကို ရွေးချယ်ထားဆဲဖြစ်နေပါလိမ့်မယ်)။ မကြာသေးခင်က သုံးခဲ့တဲ့ procedures တွေဆီကို အမြန်ပြန်သွားဖို့အတွက် toolbar ပေါ်က Dialog Recall button ကို အသုံးပြုနိုင်ပါတယ်။

Figure 28. Dialog Recall button

၂။ Charts ကို နှိပ်ပါ။

၃။ Bar charts ကို ရွေးချယ်ပြီး Continue ကို နှိပ်ပါ။

၄။ procedure ကို run ဖို့အတွက် main dialog box ထဲက OK ကို နှိပ်ပါ။

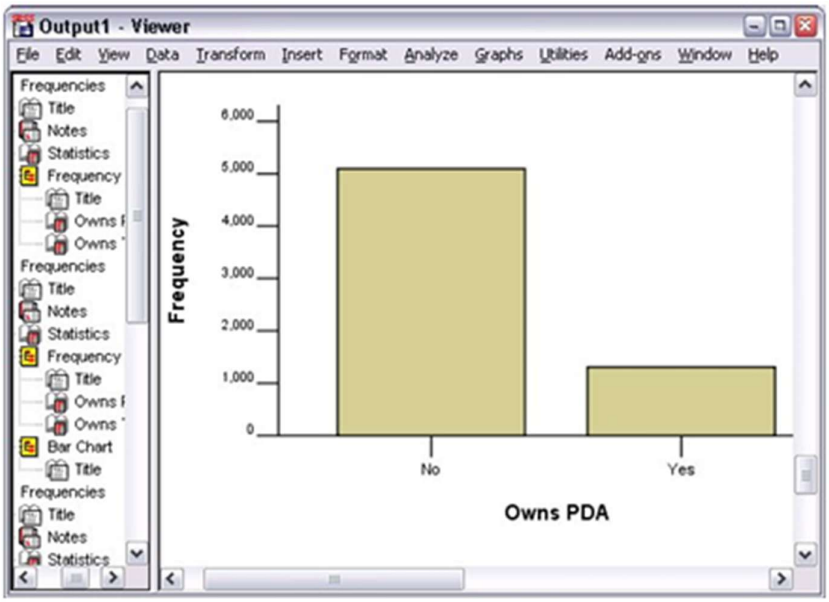


Figure 29. Bar chart

အခုဆိုရင် frequency tables တွေအပြင်၊ အဲဒီတူညီတဲ့ သတင်းအချက်အလက်တွေကို bar charts တွေပုံစံနဲ့ပါ ပြသပေးထားတဲ့အတွက် လူအများစုက PDAs တွေ မပိုင်ဆိုင်ကြပေမယ့် လူတိုင်းနီးပါးက TV ပိုင်ဆိုင်ကြတယ်ဆိုတာကို အလွယ်တကူ မြင်တွေ့နိုင်မှာ ဖြစ်ပါတယ်။

Summary Measures for Scale Variables

Scale variables တွေအတွက် အသုံးပြုနိုင်တဲ့ summary measures တွေ အများကြီးရှိပါတယ်-

- **Measures of central tendency**. အသုံးအများဆုံး measures of central tendency တွေကတော့ mean (arithmetic average လို့လည်းခေါ်တဲ့ ပျမ်းမျှကိန်း) နဲ့ median (cases တွေရဲ့ တစ်ဝက်က အထက်မှာရှိပြီး ကျန်တစ်ဝက်က အောက်မှာရှိနေတဲ့ value) တို့ပဲ ဖြစ်ပါတယ်။
- **Measures of dispersion**. Data ထဲမှာရှိတဲ့ variation ဒါမှမဟုတ် spread ဖြစ်နေမှု ပမာဏကို တိုင်းတာပေးတဲ့ statistics တွေဖြစ်ပြီး၊ ၎င်းတို့ထဲမှာ standard deviation, minimum နဲ့ maximum တို့ ပါဝင်ပါတယ်။

၁။ Frequencies dialog box ကို ပြန်ဖွင့်ပါ။

၂။ အရင် settings တွေကို ရှင်းလင်းဖို့ Reset ကို နှိပ်ပါ။

၃။ Household income in thousands [income] ကို ရွေးချယ်ပြီး Variable(s) list ထဲကို ရွှေ့ထည့်ပါ။

၄။ Statistics ကို နှိပ်ပါ။

၅။ Mean, Median, Std. deviation, Minimum နဲ့ Maximum တို့ကို ရွေးချယ်ပါ။

၆။ Continue ကို နှိပ်ပါ။

၇။ main dialog box ထဲက Display frequency tables ဆိုတဲ့ နေရာမှာ အမှန်ခြစ် ဖြုတ်လိုက်ပါ (Scale variables တွေမှာ data file ထဲက cases အရေအတွက်နီးပါးလောက် ကွဲပြားတဲ့ values တွေ အများကြီးရှိနေနိုင်တဲ့အတွက် frequency tables တွေက သိပ်ပြီး အသုံးမဝင်တတ်ပါဘူး)။

၈။ procedure ကို run ဖို့အတွက် OK ကို နှိပ်ပါ။

Frequencies Statistics table ကို Viewer window ထဲမှာ ပြသပေးပါလိမ့်မယ်။

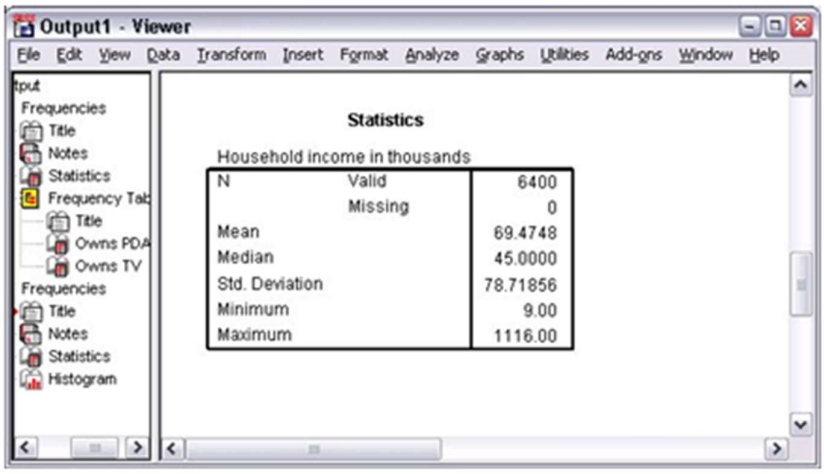


Figure 30. Frequencies Statistics table

ဒီဥပမာမှာဆိုရင် mean နဲ့ median ကြားမှာ အတော်လေး ကြီးမားတဲ့ ကွာခြားချက် ရှိနေတာကို တွေ့ရပါမယ်။ Mean ဟာ median ထက် 25,000 နီးပါးလောက် ပိုများနေပြီး၊ ဒါဟာ values တွေက normally distributed မဖြစ်ဘူးဆိုတာကို ညွှန်ပြနေတာပါ။ ဒီ distribution ကို histogram တစ်ခုသုံးပြီး မျက်မြင် စစ်ဆေးကြည့်နိုင်ပါတယ်။

Histograms for Scale Variables

၁။ Frequencies dialog box ကို ပြန်ဖွင့်ပါ။

၂။ Charts ကို နှိပ်ပါ။

၃။ Histograms နဲ့ With normal curve ကို ရွေးချယ်ပါ။

၄။ Continue ကို နှိပ်ပါ။ ပြီးရင် procedure ကို run ဖို့ main dialog box ထဲက OK ကို နှိပ်ပါ။

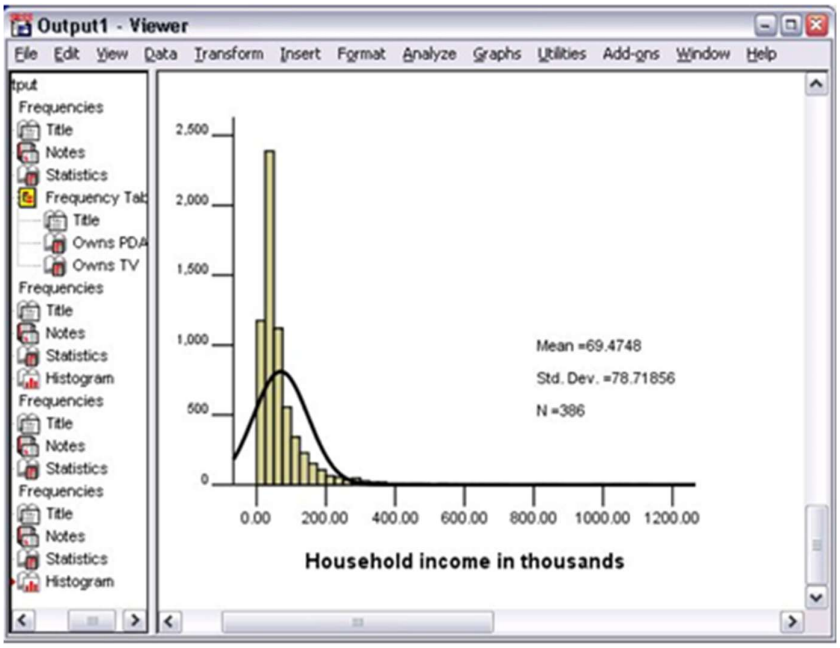


Figure 31. Histogram

အများစုသော cases တွေက scale ရဲ့ အောက်ခြေပိုင်းမှာ စုပြုံနေကြပြီး 100,000 အောက်မှာပဲ ရှိနေကြတာ တွေ့ရပါမယ်။ ဒါပေမယ့် 500,000 range နဲ့ အဲဒီအထက်မှာလည်း cases အနည်းငယ် ရှိနေပါသေးတယ် (အရမ်းနည်းလွန်းလို့ histogram ကို မပြင်ဆင်ဘဲနဲ့ဆိုရင် မြင်ရဖို့တောင် ခက်ခဲနိုင်ပါတယ်)။ အဲဒီ cases အနည်းငယ်ရဲ့ မြင့်မားတဲ့ values တွေက mean အပေါ်ကို သိသိသာသာ သက်ရောက်မှု ရှိစေပေမယ့်၊ median အပေါ်ကိုတော့ သက်ရောက်မှု အနည်းငယ် ဒါမှမဟုတ် လုံးဝကို မရှိစေပါဘူး။ ဒါကြောင့် ဒီဥပမာမှာဆိုရင် median ဟာ central tendency အတွက် ပိုကောင်းတဲ့ indicator တစ်ခု ဖြစ်နေတာပါ။

Chapter 5. Creating and editing charts

သင်ဟာ chart အမျိုးအစား (chart types) များစုံကို ဖန်တီး (create) လုပ်နိုင်ပြီး ပြင်ဆင် (edit) နိုင်ပါတယ်။ ဒီ chapter မှာတော့ bar charts တွေကို ဘယ်လို ဖန်တီးမလဲ၊ ဘယ်လို ပြင်ဆင်မလဲ ဆိုတာကို လေ့လာသွားမှာပါ။ ဒီ အခြေခံသဘောတရားတွေကို တခြား ဘယ် chart type မှာမဆို အသုံးပြုနိုင်ပါတယ်။

Chart creation basics

Chart ဖန်တီးခြင်းရဲ့ အခြေခံတွေကို ပြသဖို့အတွက်၊ job satisfaction ရဲ့ မတူညီတဲ့ levels တွေအပေါ် မူတည်ပြီး mean income ကို ပြသပေးမယ့် bar chart တစ်ခုကို ဖန်တီးကြည့်ကြပါမယ်။ ဒီဥပမာအတွက် demo.sav ဆိုတဲ့ data file ကို သုံးသွားမှာပါ (ပိုပြီး သိချင်တယ်ဆိုရင် Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ သွားကြည့်နိုင်ပါတယ်) ။

၁။ menus တွေထဲကနေ Graphs > Chart Builder... ကို ရွေးချယ်ပါ။

Chart Builder dialog box ဆိုတာက သင် chart တစ်ခုကို တည်ဆောက်နေစဉ်မှာ အဲဒီ chart က ဘယ်လိုပုံစံ ထွက်လာမလဲ ဆိုတာကို ကြိုတင် (preview) ကြည့်ရှုနိုင်တဲ့ interactive window တစ်ခု ဖြစ်ပါတယ်။

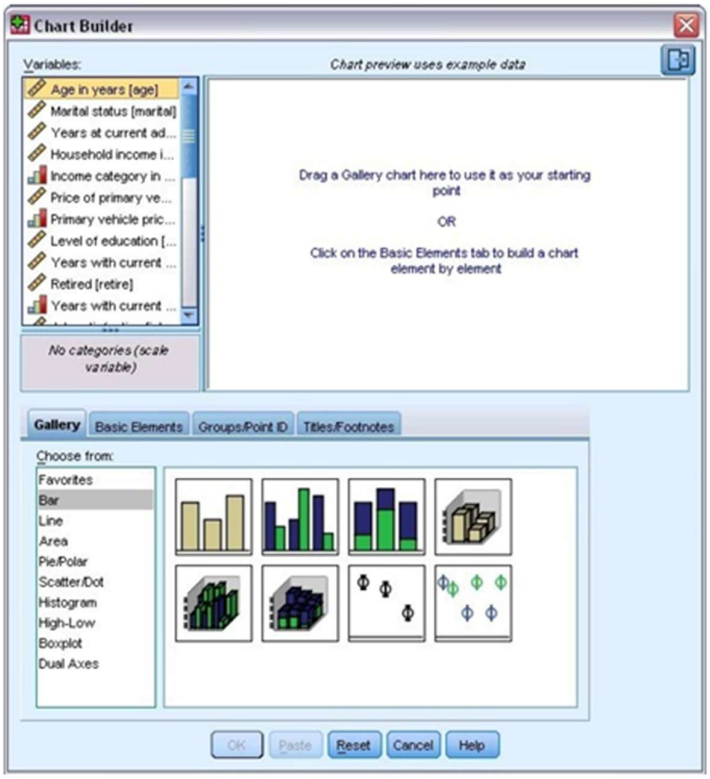


Figure 32. Chart Builder dialog box

Using the Chart Builder gallery

၁။ Gallery tab ကို ရွေးချယ်ထားခြင်း မရှိသေးဘူးဆိုရင် ၎င်းကို နှိပ်ပါ။ Gallery ထဲမှာ ကြိုတင်သတ်မှတ်ထားတဲ့ (predefined) charts တွေ အများကြီး ပါဝင်ပြီး၊ ၎င်းတို့ကို chart type အလိုက် စုစည်းထားပေးပါတယ်။ Basic Elements tab ကနေလည်း charts တွေကို အစကနေ ကိုယ်တိုင် ဖန်တီးဖို့အတွက် လိုအပ်တဲ့ အခြေခံ elements တွေ (ဥပမာ- axes တွေနဲ့ graphic elements တွေ) ကို ထောက်ပံ့ပေးထားပေမယ့်၊ Gallery ကို အသုံးပြုရတာက ပိုပြီး လွယ်ကူပါတယ်။

၂။ Bar ကို ရွေးချယ်ထားခြင်း မရှိသေးဘူးဆိုရင် ၎င်းကို နှိပ်ပါ။ Gallery ထဲမှာ ရရှိနိုင်တဲ့ bar charts တွေကို ကိုယ်စားပြုတဲ့ icons လေးတွေ dialog box ထဲမှာ ပေါ်လာပါလိမ့်မယ်။ အဲဒီ ပုံလေးတွေက တိကျတဲ့ chart type ကို ခွဲခြားသိမြင်ဖို့အတွက် လုံလောက်တဲ့ သတင်းအချက်အလက်ကို ပေးစွမ်းနိုင်ပါတယ်။ တကယ်လို့ ပိုပြီး သိချင်သေးတယ် ဆိုရင်လည်း သင့်ရဲ့ mouse cursor ကို icon လေးပေါ်မှာ ခဏတင်ထားလိုက်ရင် chart အကြောင်း ရှင်းပြပေးတဲ့ ToolTip description လေး ပေါ်လာမှာ ဖြစ်ပါတယ်။

၃။ Simple bar chart အတွက် icon လေးကို Gallery ရဲ့ အပေါ်ဘက်က ဧရိယာအကျယ်ကြီးဖြစ်တဲ့ "canvas" ပေါ်ကို ဆွဲတင် (drag and drop) လိုက်ပါ။ Chart Builder က ဖန်တီးလိုက်တဲ့ chart ရဲ့ preview ကို canvas ပေါ်မှာ ပြသပေးပါလိမ့်မယ်။ ဒီ chart ကို ရေးဆွဲဖို့ အသုံးပြုထားတဲ့ data တွေက သင့်ရဲ့ အမှန်တကယ် (actual) data တွေ မဟုတ်ဘဲ example data တွေသာ ဖြစ်တယ်ဆိုတာကို သတိပြုပါ။

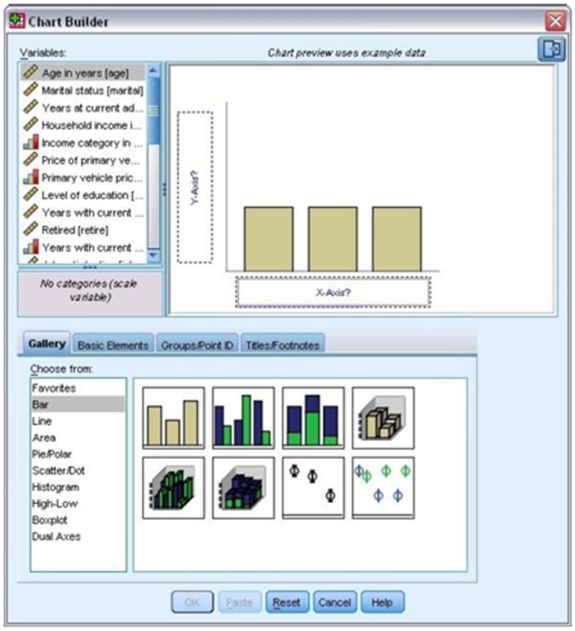


Figure 33. Bar chart on Chart Builder canvas

Defining variables and statistics

Canvas ပေါ်မှာ chart တစ်ခု ရှိနေပြီ ဆိုပေမယ့်၊ ဘား (bars) တွေရဲ့ အမြင့်ကို ဘယ်လိုထားမလဲ၊ ဘယ် variable category ကို ဘယ် bar က ကိုယ်စားပြုမလဲ ဆိုတာတွေကို သတ်မှတ်ပေးဖို့အတွက် variables တွေ ဒါမှမဟုတ် statistics တွေ မရှိသေးတဲ့အတွက် ဒါက ပြည့်စုံတဲ့ chart တစ်ခု မဟုတ်သေးပါဘူး။ Variables တွေနဲ့ statistics တွေ မပါဘဲနဲ့ chart တစ်ခုကို ဖန်တီးလို့ မရပါဘူး။ Variables တွေကို ထည့်သွင်းဖို့အတွက် canvas ရဲ့ ဘယ်ဘက်မှာရှိတဲ့ Variables list ထဲကနေ ဆွဲယူ (drag) ထည့်သွင်းပေးရမှာ ဖြစ်ပါတယ်။

Chart Builder မှာ variable တစ်ခုရဲ့ measurement level က အရမ်းကို အရေးကြီးပါတယ်။ သင်က Job satisfaction variable ကို x axis ပေါ်မှာ အသုံးပြုမှာပါ။ ဒါပေမယ့် variable ရဲ့ ဘေးနားက (ပေတံပုံစံလေးနဲ့) icon က ၎င်းရဲ့ measurement level ဟာ scale အနေနဲ့ သတ်မှတ်ထားကြောင်း ညွှန်ပြနေပါတယ်။ မှန်ကန်တဲ့ chart တစ်ခုကို ဖန်တီးဖို့အတွက် ဆိုရင်တော့ သင်ဟာ categorical measurement level ကို အသုံးပြုဖို့ လိုအပ်ပါတယ်။ Measurement level ကို ပြောင်းလဲဖို့ Variable View ဆီကို ပြန်သွားနေမယ့်အစား၊ Chart Builder ထဲမှာတင် measurement level ကို ယာယီ ပြောင်းလဲလို့ ရပါတယ်။

၁။ Variables list ထဲက Job satisfaction ကို right-click နှိပ်ပြီး Ordinal ကို ရွေးချယ်ပါ။ Job satisfaction ထဲက categories တွေကို ကျေနပ်မှုအဆင့် (level of satisfaction) အလိုက် အစဉ်လိုက် (ranked) သတ်မှတ်နိုင်တဲ့အတွက် Ordinal က သင့်လျော်တဲ့ measurement level တစ်ခု ဖြစ်ပါတယ်။ Measurement level ကို ပြောင်းလဲလိုက်ပြီးတဲ့နောက်မှာ icon လေးလည်း ပြောင်းလဲသွားတာကို သတိပြုပါ။

၂။ အခု Job satisfaction ကို Variables list ထဲကနေ x axis drop zone ဆီကို ဆွဲထည့် (drag) လိုက်ပါ။

y axis drop zone ကတော့ Count statistic အနေနဲ့ default သတ်မှတ်ထားပါတယ်။ တကယ်လို့ သင်က (percentage သို့မဟုတ် mean လိုမျိုး) တခြား statistic တစ်ခုကို အသုံးပြုချင်တယ်ဆိုရင် အလွယ်တကူ ပြောင်းလဲနိုင်ပါတယ်။ ဒီဥပမာမှာတော့ အဲဒီ statistics နှစ်ခုလုံးကို အသုံးပြုသွားမှာ မဟုတ်ပေမယ့်၊ နောက်တစ်ချိန် သင် statistic ကို ပြောင်းလဲချင်တဲ့အခါ အထောက်အကူဖြစ်စေဖို့အတွက် အဲဒီလုပ်ငန်းစဉ်ကို ပြန်လည် ဆွေးနွေးကြည့်ပါမယ်။

၃။ Chart Builder ရဲ့ ဘေးဘက် (side bar) မှာရှိတဲ့ Element Properties tab ကို နှိပ်ပါ။ (တကယ်လို့ side bar ကို မမြင်ရဘူးဆိုရင် side bar ပေါ်လာအောင် Chart Builder ရဲ့ ညာဘက်အပေါ်ထောင့်မှာရှိတဲ့ button လေးကို နှိပ်ပါ)။

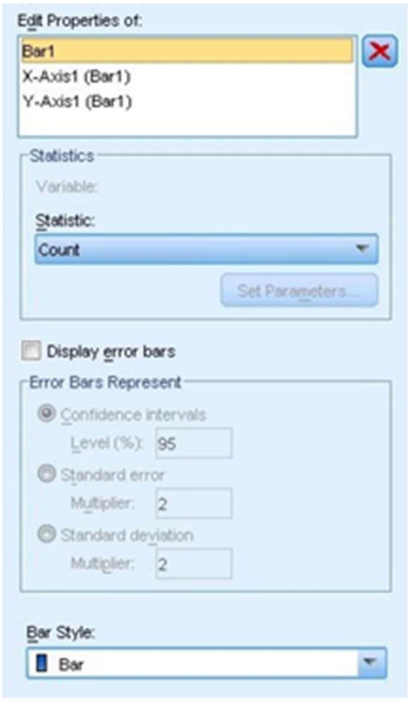


Figure 34. Element Properties

Element Properties ဟာ မတူညီတဲ့ chart elements တွေရဲ့ properties တွေကို ပြောင်းလဲနိုင်အောင် ကူညီပေးပါတယ်။ ဒီ elements တွေထဲမှာ (bar chart ထဲက bars တွေလိုမျိုး) graphic elements တွေနဲ့ chart ပေါ်က axes တွေ ပါဝင်ပါတယ်။ သတ်မှတ်ထားတဲ့ element တစ်ခုရဲ့ properties ကို ပြောင်းလဲဖို့အတွက် Edit Properties of list ထဲကနေ သင်ပြောင်းချင်တဲ့ element ကို ရွေးချယ်ပါ။ List ရဲ့ ညာဘက်မှာရှိတဲ့ အနီရောင် X လေးကိုလည်း သတိပြုပါ။ ဒီ button က canvas ပေါ်ကနေ graphic element တစ်ခုကို ဖျက်ပစ် (delete) ဖို့ အသုံးပြုတာပါ။ Bar1 ကို ရွေးချယ်ထားတဲ့အတွက် ပြသထားတဲ့ properties တွေက graphic elements တွေ၊ အထူးသဖြင့် bar graphic element နဲ့ သက်ဆိုင်ပါတယ်။

Statistic drop-down list ကနေ အသုံးပြုနိုင်တဲ့ တိကျတဲ့ statistics တွေကို ပြသပေးပါတယ်။ Chart type တိုင်းနီးပါးအတွက် ဒီတူညီတဲ့ statistics တွေကို အသုံးပြုနိုင်ပါတယ်။ တချို့ statistics တွေဟာ y axis drop zone ထဲမှာ variable တစ်ခု ရှိနေဖို့ လိုအပ်တယ်ဆိုတာကိုတော့ သတိပြုပါ။

၄။ Household income in thousands ကို Variables list ထဲကနေ y axis drop zone ဆီ ဆွဲထည့်ပါ။ y axis ပေါ်က variable က scalar ဖြစ်ပြီး x axis variable က categorical (ordinal ဟာ categorical measurement level ရဲ့ အမျိုးအစားတစ်ခု ဖြစ်ပါတယ်) ဖြစ်နေတဲ့အတွက်၊ y axis drop zone က Mean statistic အနေနဲ့ default ပြောင်းသွားပါလိမ့်မယ်။ ဒါတွေက သင်အလိုရှိတဲ့

variables တွေနဲ့ statistics တွေ ဖြစ်တဲ့အတွက် element properties ကို ပြောင်းလဲဖို့ မလိုအပ်တော့ပါဘူး။

Adding text

Chart ထဲမှာ titles တွေနဲ့ footnotes တွေကိုလည်း ထည့်သွင်းနိုင်ပါတယ်။

၁။ Titles/Footnotes tab ကို နှိပ်ပါ။

၂။ Title 1 ကို ရွေးချယ်ပါ။

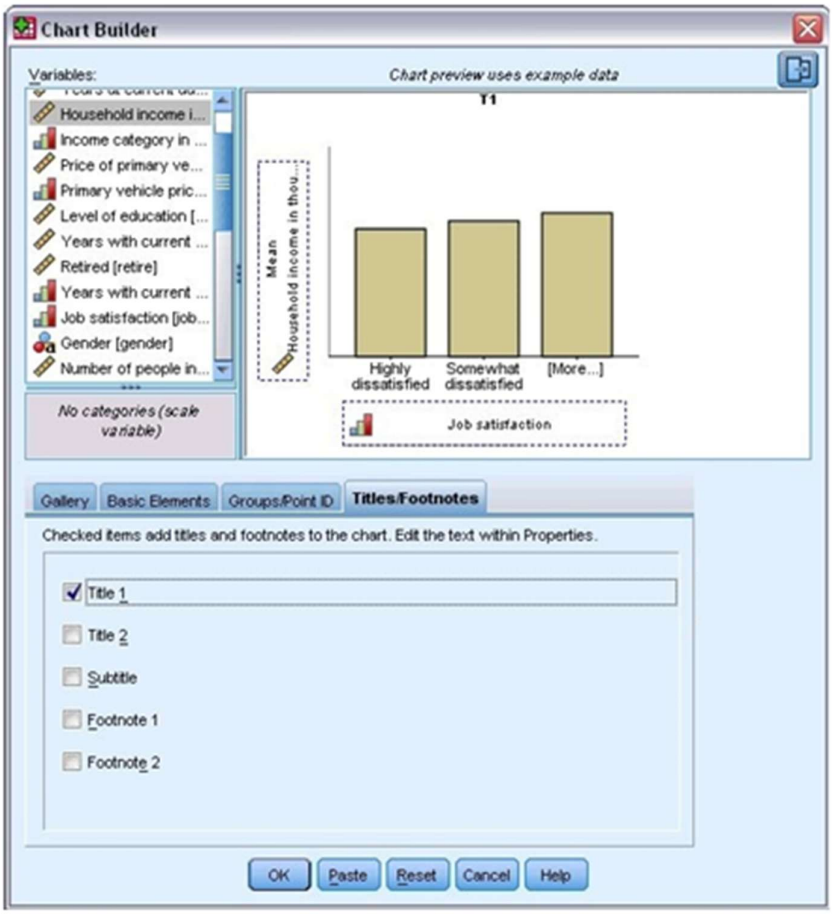


Figure 35. Title 1 displayed on canvas

Canvas ပေါ်မှာ T1 ဆိုတဲ့ label နဲ့အတူ title တစ်ခု ပေါ်လာပါလိမ့်မယ်။

၃။ Element Properties tab ထဲက Edit Properties of list မှာ Title 1 ကို ရွေးချယ်ပါ။

၄။ Content text box ထဲမှာ Income by Job Satisfaction လို့ ရိုက်ထည့်ပါ။ ဒါက title အနေနဲ့ ပြသပေးမယ့် text ဖြစ်ပါတယ်။

Creating the chart

၁။ Bar chart ကို ဖန်တီးဖို့ OK ကို နှိပ်ပါ။

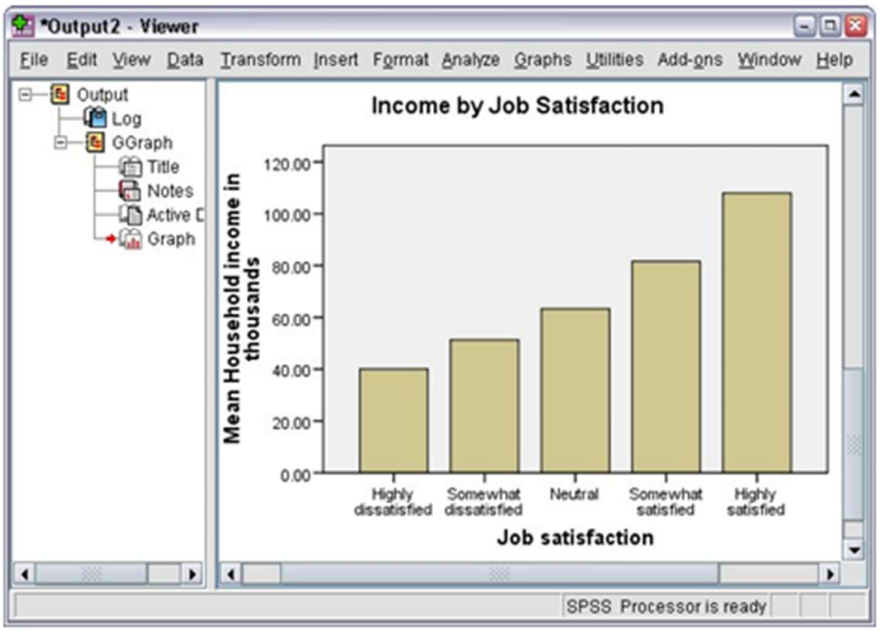


Figure 36. Bar chart

ဒီ bar chart အရ သူတို့ရဲ့ အလုပ် (jobs) တွေကို ပိုပြီး ကျေနပ်မှုရှိတဲ့ ဖြေဆိုသူ (respondents) တွေဟာ ပိုမိုမြင့်မားတဲ့ အိမ်ထောင်စု ဝင်ငွေ (household incomes) တွေ ရှိကြတယ် ဆိုတာကို ဖော်ပြနေပါတယ်။

Chapter 6. Working with Output

Statistical procedure တစ်ခုကို run လိုက်လို့ ထွက်လာတဲ့ results တွေကို Viewer ထဲမှာ ပြသပေးပါတယ်။ သင် run တဲ့ procedure အပေါ်မူတည်ပြီး ထွက်လာမယ့် output က statistical tables တွေ၊ charts တွေ၊ graphs တွေ ဒါမှမဟုတ် text တွေ ဖြစ်နိုင်ပါတယ်။ ဒီအခန်းမှာတော့ viewertut.spv နဲ့ demo.sav ဆိုတဲ့ files တွေကို သုံးသွားမှာပါ။ ပိုပြီးသိချင်တယ်ဆိုရင်တော့ Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ သွားရောက်လေ့လာနိုင်ပါတယ်။

Using the Viewer

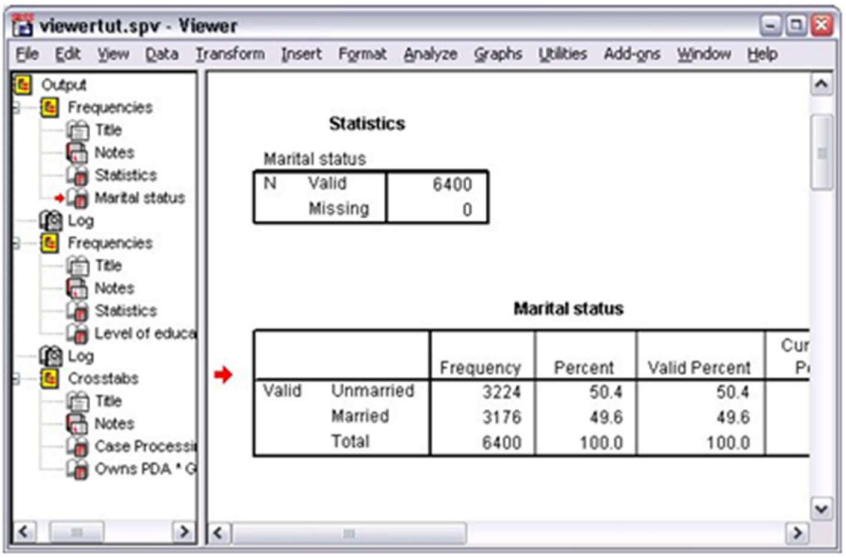


Figure 37. Viewer

Viewer window ကို panes (အကန့်) နှစ်ခု ခွဲခြားထားပါတယ်။ **Outline pane** ထဲမှာ Viewer ထဲမှာ သိမ်းဆည်းထားတဲ့ သတင်းအချက်အလက်အားလုံးရဲ့ အကျဉ်းချုပ် (outline) ပါဝင်ပါတယ်။ **Contents pane** ထဲမှာတော့ statistical tables တွေ၊ charts တွေနဲ့ text output တွေ ပါဝင်ပါတယ်။

Window ရဲ့ contents တွေထဲမှာ အပေါ်အောက်၊ ဘယ်ညာ ကြည့်ဖို့အတွက် scroll bars တွေကို သုံးနိုင်ပါတယ်။ ပိုပြီးလွယ်လွယ်ကူကူ ကြည့်နိုင်ဖို့အတွက် outline pane ထဲက item တစ်ခုကို နှိပ်လိုက်ရင် contents pane ထဲမှာ အဲဒီ item ကို ပြသပေးမှာပါ။

၁။ Outline pane ရဲ့ အကျယ်ကို ပြောင်းလဲဖို့ ညာဘက်အစွန်း (right border) ကို ကလစ်နှိပ်ပြီး ဆွဲ (drag) လို့ရပါတယ်။

Outline pane ထဲက ဖွင့်ထားတဲ့ စာအုပ်ပုံ (open book icon) လေးက အဲဒီအရာကို Viewer မှာ လက်ရှိ မြင်တွေ့နေရတယ်လို့ ဆိုလိုတာပါ (contents pane ရဲ့ မြင်ရတဲ့အပိုင်းမှာ လောလောဆယ် မရှိနေရင်တောင်မှပေါ့) ။

၂။ Table ဒါမှမဟုတ် chart တစ်ခုကို ဖျောက် (hide) ထားချင်တယ်ဆိုရင်၊ outline pane ထဲမှာရှိတဲ့ သူ့ရဲ့ book icon လေးကို double-click နှိပ်ပါ။

အဲဒီအခါ ပိတ်ထားတဲ့ စာအုပ်ပုံ (closed book icon) လေးအဖြစ် ပြောင်းသွားပြီး၊ ၎င်းနဲ့သက်ဆိုင်တဲ့ သတင်းအချက်အလက်တွေကို အခု ဖျောက်ထားပြီဖြစ်ကြောင်း ညွှန်ပြနေမှာပါ။

၃။ ဖျောက်ထားတဲ့ output ကို ပြန်ပေါ်ချင်တယ်ဆိုရင် closed book icon လေးကို double-click ပြန်နှိပ်ပါ။

သတ်မှတ်ထားတဲ့ statistical procedure တစ်ခုကနေ ထွက်လာတဲ့ output အားလုံးကို ဖြစ်စေ၊ Viewer ထဲက output အားလုံးကို ဖြစ်စေ ဖျောက်ထားလို့ ရပါသေးတယ်။

၄။ သင်ဖျောက်ထားချင်တဲ့ results တွေရဲ့ procedure ဘယ်ဘက်မှာရှိတဲ့ အနှုတ်လက္ခဏာ (–) လေးပါတဲ့ box ကို နှိပ်ပါ၊ ဒါမှမဟုတ် output အားလုံးကို ဖျောက်ချင်ရင်တော့ outline pane ရဲ့ အပေါ်ဆုံး item ဘေးက box ကို နှိပ်ပါ။ Outline လေးက ခေါက်ဝင်သွားပြီး၊ ဒီ results တွေကို ဖျောက်ထားကြောင်း မြင်သာအောင် ပြသပေးပါလိမ့်မယ်။ Output ပြသတဲ့ အစီအစဉ် (order) ကိုလည်း ပြောင်းလဲနိုင်ပါသေးတယ်။

၅။ Outline pane ထဲမှာ သင်ရွှေ့ချင်တဲ့ items တွေကို ကလစ်နှိပ်ပါ။

၆။ ရွှေးချယ်ထားတဲ့ items တွေကို outline ထဲက နေရာအသစ်ဆီ ဆွဲထည့် (drag) လိုက်ပါ။

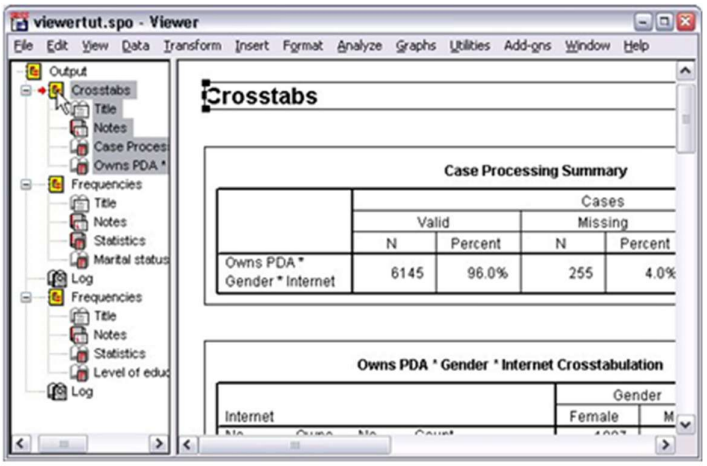


Figure 38. Reordered output in the Viewer

Contents pane ထဲမှာလည်း output items တွေကို ကလစ်နှိပ်ပြီး ဆွဲရွှေ့ခြင်းဖြင့် နေရာပြောင်းလို့ ရပါတယ်။

Using the Pivot Table Editor

Statistical procedures အများစုကနေ ထွက်လာတဲ့ results တွေကို pivot tables တွေအနေနဲ့ ပြသပေးပါတယ်။

Accessing Output Definitions

Statistical terms တော်တော်များများကို output ထဲမှာ ပြသပေးထားပါတယ်။ ဒီ terms တွေရဲ့ အဓိပ္ပာယ် ဖွင့်ဆိုချက်တွေကို Viewer ထဲမှာတင် တိုက်ရိုက် ကြည့်ရှုနိုင်ပါတယ်။

၁။ Owns PDA * Gender * Internet Crosstabulation table ကို double-click နှိပ်ပါ။

၂။ Expected Count ကို right-click နှိပ်ပြီး pop-up menu ထဲက What's This? ကို ရွေးချယ်ပါ။

အဓိပ္ပာယ် ဖွင့်ဆိုချက်ကို pop-up window လေးတစ်ခုထဲမှာ ပြသပေးပါလိမ့်မယ်။

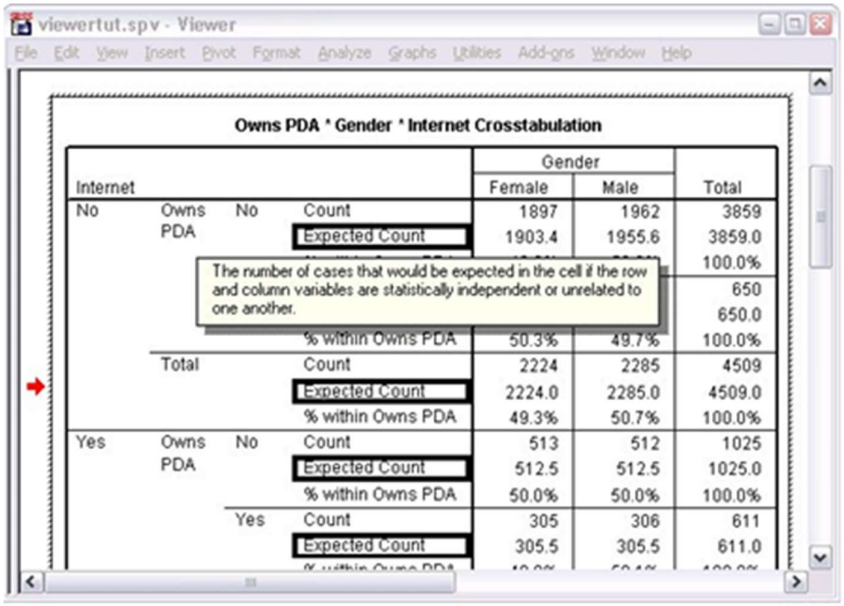


Figure 39. Pop-up definition

Pivoting Tables

ပုံမှန်ထုတ်ပေးလိုက်တဲ့ (default) tables တွေက သင်လိုချင်သလောက် သပ်ရပ်ရှင်းလင်းစွာ ပြသနိုင်ချင်မှ ပြသနိုင်ပါလိမ့်မယ်။ Pivot tables တွေကို သုံးပြီး rows တွေနဲ့ columns တွေကို နေရာချင်း ပြောင်း (transpose ဒါမှမဟုတ် "flip" လုပ်) တာ၊ table ထဲက data အစီအစဉ်ကို ချိန်ညှိတာတွေအပြင် တခြားနည်းလမ်းမျိုးစုံနဲ့ပါ table ကို ပြင်ဆင် (modify) နိုင်ပါတယ်။ ဥပမာ၊ rows တွေနဲ့ columns တွေကို နေရာချင်းပြောင်းလိုက်ခြင်းအားဖြင့် တိုပြီးကျယ်တဲ့ table တစ်ခုကို ရှည်ပြီးကျဉ်းတဲ့ table တစ်ခုအဖြစ် ပြောင်းလဲနိုင်ပါတယ်။ Table ရဲ့ layout ကို

ပြောင်းလဲလိုက်ပေမယ့် results တွေအပေါ်တော့ ထိခိုက်မှု မရှိပါဘူး။ ကိုယ့်ရဲ့ သတင်းအချက်အလက်တွေကို တခြားနည်းလမ်း ဒါမှမဟုတ် ပိုကောင်းတဲ့နည်းလမ်းနဲ့ ပြသဖို့အတွက် လုပ်ဆောင်တာပဲ ဖြစ်ပါတယ်။

၁။ အကယ်၍ activate မဖြစ်သေးဘူးဆိုရင် Owns PDA * Gender * Internet Crosstabulation table ကို activate လုပ်ဖို့ သူ့ကို double-click နှိပ်ပါ။

၂။ Pivoting Trays window က မပေါ်နေဘူးဆိုရင်၊ menus တွေထဲကနေ Pivot > Pivoting Trays ကို ရွေးချယ်ပါ။

Pivoting trays တွေက data တွေကို columns တွေ၊ rows တွေနဲ့ layers တွေကြားမှာ ရွှေ့ပြောင်းနိုင်တဲ့ နည်းလမ်းတစ်ခု ဖြစ်ပါတယ်။

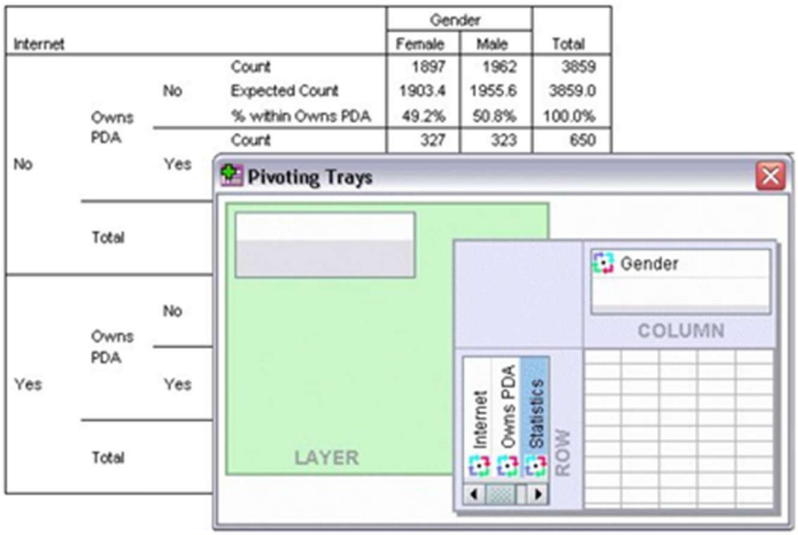


Figure 40. Pivoting trays

၃။ Statistics element ကို Row dimension ကနေ Column dimension ထဲက Gender အောက်ဆီ ဆွဲထည့်ပါ။ Table က သင်ပြောင်းလဲလိုက်တဲ့အတိုင်း ချက်ချင်း ပြောင်းလဲသွားပါလိမ့်မယ်။ Pivoting tray ထဲက elements တွေရဲ့ အစီအစဉ်က table ထဲမှာရှိတဲ့ elements တွေရဲ့ အစီအစဉ်အတိုင်းပါပဲ။

၄။ ဒီ rows နှစ်ခုရဲ့ အစီအစဉ်ကို ပြောင်းပြန်လုပ်ဖို့အတွက် row dimension ထဲမှာ Owns PDA element ကို Internet element ရဲ့အရှေ့ကို ဆွဲယူ (drag and drop) ထည့်လိုက်ပါ။

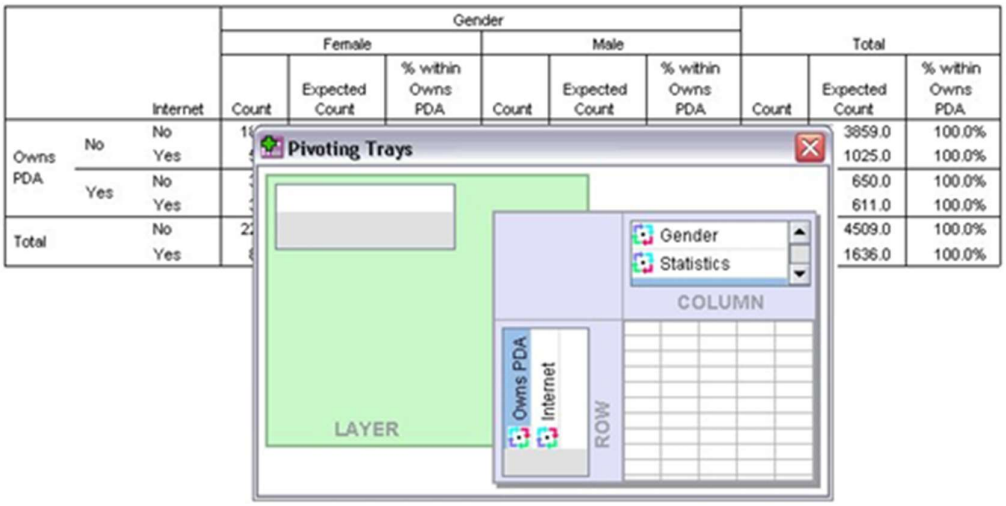


Figure 41. Swap rows

Creating and Displaying Layers

Nested categories အချက်အလက်တွေ ပါဝင်တဲ့ ကြီးမားတဲ့ tables တွေအတွက် layers တွေ ဖန်တီးတာက အသုံးဝင်ပါတယ်။ Layers တွေ ဖန်တီးလိုက်ခြင်းအားဖြင့် table ရဲ့ အသွင်အပြင်ကို ရိုးရှင်းသွားစေပြီး၊ ဖတ်ရပိုလွယ်ကူစေပါတယ်။

၁။ Gender element ကို Column dimension ကနေ Layer dimension ဆီ ဆွဲထည့်ပါ။

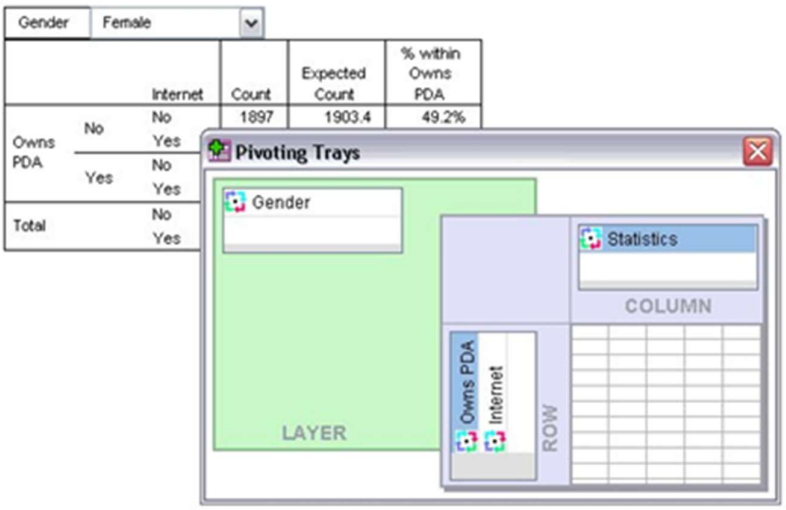


Figure 42. Gender pivot icon in the Layer dimension

တခြား layer တစ်ခုကို ပြသချင်တယ်ဆိုရင်၊ table ထဲက drop-down list ကနေ category တစ်ခုကို ရွေးချယ်နိုင်ပါတယ်။

Editing Tables

သင်က custom TableLook ကို အချိန်ယူပြီး မဖန်တီးထားဘူးဆိုရင်၊ pivot tables တွေကို standard formatting နဲ့ပဲ ဖန်တီးပေးမှာပါ။ Table တစ်ခုအတွင်းမှာရှိတဲ့ ဘယ် text ရဲ့ formatting ကိုမဆို ပြောင်းလဲနိုင်ပါတယ်။ ပြောင်းလဲနိုင်တဲ့ formats တွေထဲမှာ font name, font size, font style (bold ဒါမှမဟုတ် italic) နဲ့ color တို့ ပါဝင်ပါတယ်။

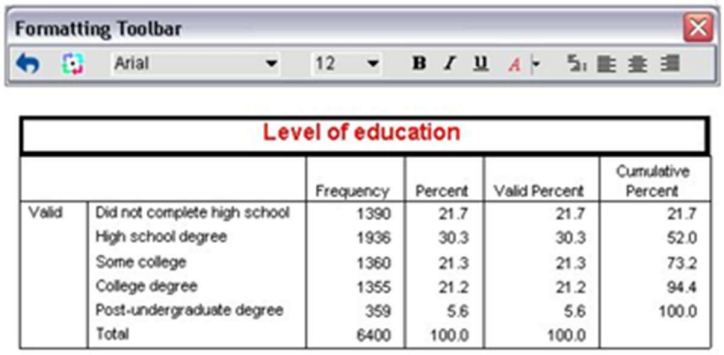
၁။ Level of education table ကို double-click နှိပ်ပါ။

၂။ Formatting toolbar က မပေါ်နေဘူးဆိုရင်၊ menus တွေထဲကနေ View > Toolbar ကို ရွေးချယ်ပါ။

၃။ Level of education ဆိုတဲ့ title text ကို ကလစ်နှိပ်ပါ။

၄။ Toolbar ပေါ်က font sizes drop-down list ထဲကနေ 12 ကို ရွေးချယ်ပါ။

၅။ Title text ရဲ့ အရောင်ကို ပြောင်းဖို့အတွက် text color tool ကို နှိပ်ပြီး အရောင်အသစ်တစ်ခုကို ရွေးချယ်ပါ။



Level of education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Did not complete high school	1390	21.7	21.7	21.7
	High school degree	1936	30.3	30.3	52.0
	Some college	1360	21.3	21.3	73.2
	College degree	1355	21.2	21.2	94.4
	Post-undergraduate degree	359	5.6	5.6	100.0
	Total	6400	100.0	100.0	

Figure 43. Reformatted title text in the pivot table

Tables တွေနဲ့ labels တွေရဲ့ အကြောင်းအရာ (contents) တွေကိုလည်း edit လုပ်လို့ရပါသေးတယ်။ ဥပမာအနေနဲ့ ဒီ table ရဲ့ title ကို ပြောင်းလဲနိုင်ပါတယ်။

၆။ Title ကို double-click နှိပ်ပါ။

၇။ Label အသစ်အနေနဲ့ Education Level လို့ ရိုက်ထည့်ပါ။ မှတ်ချက် - အကယ်၍ သင်က table ထဲက values တွေကို ပြောင်းလဲလိုက်မယ်ဆိုရင်၊ totals တွေနဲ့ တခြား statistics တွေကို ပြန်တွက်ပေး (recalculated) မှာ မဟုတ်ပါဘူး။

Hiding Rows and Columns

Table ထဲမှာ ပြသထားတဲ့ data တချို့က အသုံးမဝင်တာမျိုး၊ ဒါမှမဟုတ် table ကို မလိုအပ်ဘဲ ရှုပ်ထွေးသွားစေတာမျိုး ဖြစ်နိုင်ပါတယ်။ ကံကောင်းတာက data တွေ မဆုံးရှုံးစေဘဲ rows တွေ columns တွေ တစ်ခုလုံးကို ဖျောက် (hide) ထားလို့ ရပါတယ်။

၁။ အကယ်၍ activate မဖြစ်သေးဘူးဆိုရင် Education Level table ကို activate လုပ်ဖို့ သူ့ကို double-click နှိပ်ပါ။

၂။ Valid Percent column label ကို ရွေးချယ်ဖို့ ကလစ်နှိပ်ပါ။

၃။ Edit menu ကနေဖြစ်ဖြစ်၊ right-click နှိပ်လို့ပေါ်လာတဲ့ pop-up menu ကနေဖြစ်ဖြစ် Select > Data and Label Cells ကို ရွေးချယ်ပါ။

၄။ View menu ကနေဖြစ်ဖြစ်၊ right-click pop-up menu ကနေဖြစ်ဖြစ် Hide (ဒါမှမဟုတ် Hide Category) ကို ရွေးချယ်ပါ။ အခုဆိုရင် အဲဒီ column ဟာ ဖျောက်ထားခံရပေမယ့် delete လုပ်လိုက်တာတော့ မဟုတ်ပါဘူး။

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	21.7	21.7
	High school degree	1936	30.3	52.0
	Some college	1360	21.3	73.2
	College degree	1355	21.2	94.4
	Post-undergraduate degree	359	5.6	100.0
	Total	6400	100.0	

Figure 44. Valid Percent column hidden in table

Column ကို ပြန်ပေါ်ချင်ရင်တော့-

၅။ menus တွေထဲကနေ View > Show All ကို ရွေးချယ်ပါ။ Rows တွေကိုလည်း columns တွေနည်းတူ ဖျောက်ထားလို့၊ ပြန်ပေါ်အောင်လုပ်လို့ ရပါတယ်။

Changing Data Display Formats

Pivot tables တွေထဲက data ပြသတဲ့ format ကိုလည်း အလွယ်တကူ ပြောင်းလဲနိုင်ပါတယ်။

၁။ အကယ်၍ activate မဖြစ်သေးဘူးဆိုရင် Education Level table ကို activate လုပ်ဖို့ သူ့ကို double-click နှိပ်ပါ။

၂။ Percent column label ကို ရွေးချယ်ဖို့ ကလစ်နှိပ်ပါ။

၃။ Edit menu ဒါမှမဟုတ် right-click pop-up menu ထဲကနေ Select > Data Cells ကို ရွေးချယ်ပါ။

၄။ Format menu ဒါမှမဟုတ် right-click pop-up menu ထဲကနေ Cell Properties ကို ရွေးချယ်ပါ။

၅။ Format Value tab ကို နှိပ်ပါ။

၆။ ဒီ column မှာ ဒသမကိန်း (decimal points) အားလုံးကို ဖျောက်ဖို့ Decimals field ထဲမှာ 0 လို့ ရိုက်ထည့်ပါ။

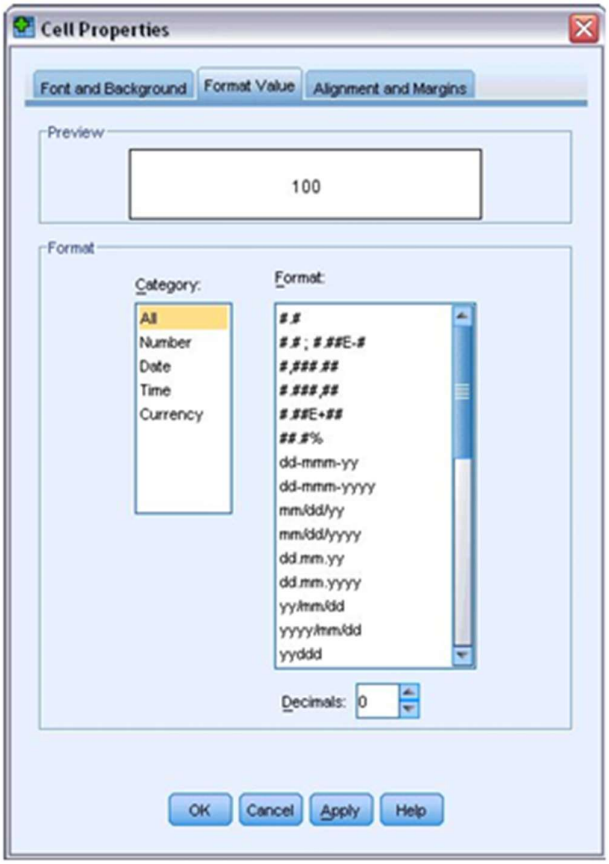


Figure 45. Cell Properties, Format Value tab

ဒီ dialog box ထဲမှာ data type နဲ့ format ကိုလည်း ပြောင်းလဲနိုင်ပါသေးတယ်။

၇။ Category list ထဲကနေ သင်လိုချင်တဲ့ type ကို ရွေးချယ်ပြီး၊ အဲဒီ type အတွက် format ကို Format list ထဲမှာ ရွေးချယ်ပါ။

၈။ ပြောင်းလဲမှုတွေကို အသက်ဝင်စေဖို့ OK ဒါမှမဟုတ် Apply ကို နှိပ်ပါ။

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	22	21.7
	High school degree	1936	30	52.0
	Some college	1360	21	73.2
	College degree	1355	21	94.4
	Post-undergraduate degree	359	6	100.0
	Total	6400	100	

Figure 46. Decimals hidden in Percent column

အခုဆိုရင် Percent column မှာ ဒသမကိန်းတွေကို ဖျောက်ထားပြီး ဖြစ်ပါတယ်။

TableLooks

Tables တွေရဲ့ format ဟာ ရှင်းလင်းပြတ်သားပြီး အဓိပ္ပာယ်ပြည့်ဝတဲ့ results တွေပေးနိုင်ဖို့ အလွန်အရေးကြီးတဲ့ အစိတ်အပိုင်းတစ်ခုပါ။ တကယ်လို့ သင့်ရဲ့ table က ဖတ်ရခက်နေမယ်ဆိုရင်၊ အဲဒီ table ထဲမှာပါတဲ့ သတင်းအချက်အလက်တွေကို အလွယ်တကူ နားလည်နိုင်မှာ မဟုတ်ပါဘူး။

Using Predefined Formats

၁။ Marital status table ကို double-click နှိပ်ပါ။

၂။ menus တွေထဲကနေ Format > TableLooks... ကို ရွေးချယ်ပါ။

TableLooks dialog box မှာ ကြိုတင်သတ်မှတ်ထားတဲ့ styles (predefined styles) မျိုးစုံကို စာရင်းနဲ့ ပြသပေးထားပါတယ်။ အဲဒီ list ထဲကနေ style တစ်ခုကို ရွေးချယ်ပြီး ညာဘက်မှာရှိတဲ့ Sample window ထဲမှာ preview ကြည့်နိုင်ပါတယ်။

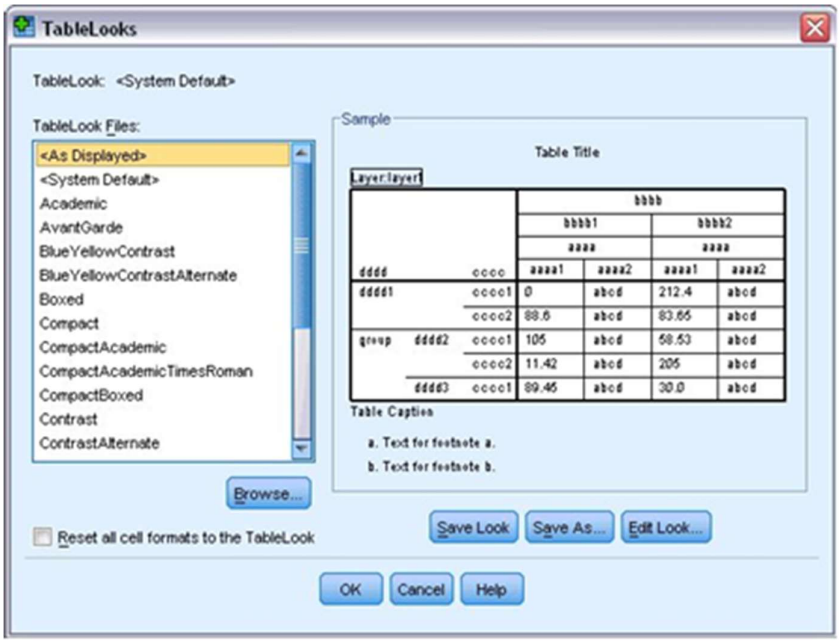


Figure 47. TableLooks dialog box

Style တစ်ခုကို ဒီအတိုင်းပဲဖြစ်ဖြစ် သုံးနိုင်သလို၊ သင့်လိုအပ်ချက်နဲ့ ပိုကိုက်ညီအောင် ရှိပြီးသား style တစ်ခုကို edit လုပ်ပြီးတော့လည်း သုံးနိုင်ပါတယ်။

၃။ ရှိပြီးသား style တစ်ခုကို သုံးဖို့အတွက် တစ်ခုကို ရွေးချယ်ပြီး OK နှိပ်ပါ။

Customizing TableLook Styles

သင့်ရဲ့ သီးခြားလိုအပ်ချက်တွေနဲ့ ကိုက်ညီဖို့အတွက် format တစ်ခုကို customize လုပ်နိုင်ပါတယ်။ နောက်ခံအရောင် (background color) ကနေစလို့ border styles တွေအထိ table တစ်ခုရဲ့ အစိတ်အပိုင်းအားလုံးနီးပါးကို customize လုပ်လို့ရပါတယ်။

၁။ Marital status table ကို double-click နှိပ်ပါ။

၂။ menus တွေထဲကနေ Format > TableLooks... ကို ရွေးချယ်ပါ။

၃။ သင်လိုချင်တဲ့ format နဲ့ အနီးစပ်ဆုံးတူတဲ့ style ကို ရွေးချယ်ပြီး Edit Look ကို နှိပ်ပါ။

၄။ Formatting options တွေကို ကြည့်ဖို့ Cell Formats tab ကို နှိပ်ပါ။

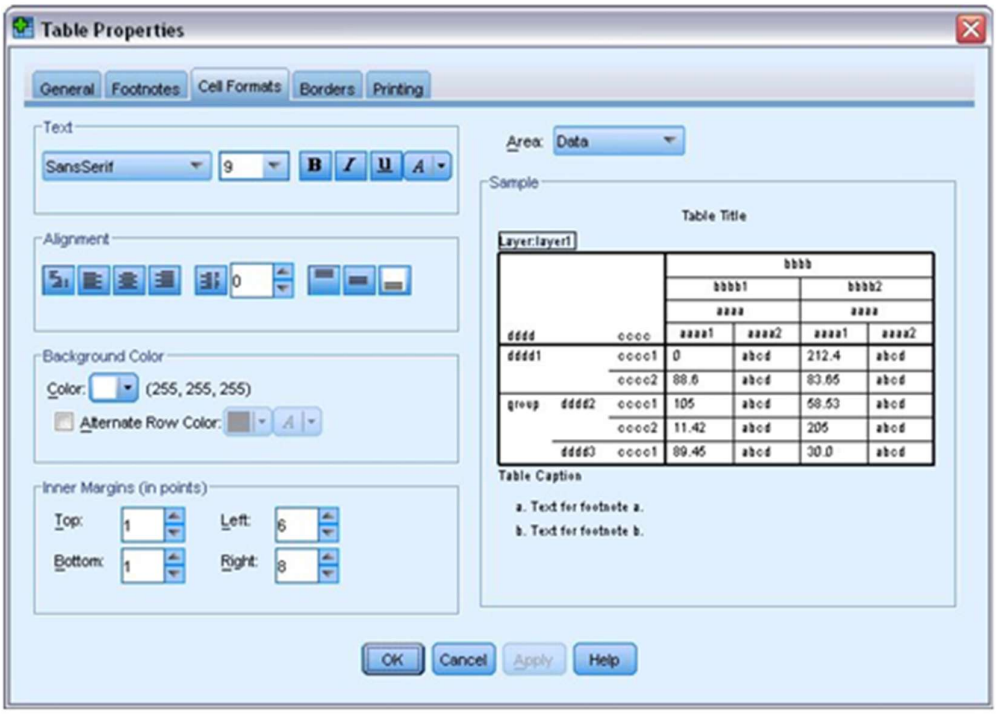


Figure 48. Table Properties dialog box

Formatting options တွေထဲမှာ font name, font size, style နဲ့ color တို့ ပါဝင်ပါတယ်။ နောက်ထပ် options တွေအနေနဲ့ alignment, text and background colors နဲ့ margin sizes တို့လည်း ပါဝင်ပါတယ်။

ဥာဘက်က Sample window မှာ သင်ပြောင်းလဲလိုက်တဲ့ formatting က table အပေါ် ဘယ်လိုသက်ရောက်မှုရှိလဲ ဆိုတာ preview ပြသပေးပါတယ်။ Table ရဲ့ ဧရိယာတစ်ခုစီအတွက် မတူညီတဲ့ formatting styles တွေ ထားရှိနိုင်ပါတယ်။ ဥပမာအားဖြင့်၊ သင်က title ကို data ရဲ့ style နဲ့ တူအောင် ထားချင်မှာ မဟုတ်ပါဘူး။ Edit လုပ်မယ့် table ဧရိယာကို ရွေးချယ်ဖို့အတွက်၊ Area drop-down list ထဲကနေ နာမည်နဲ့ရွေးချယ်လို့ရသလို၊ Sample window ထဲမှာ သင်ပြောင်းချင်တဲ့ ဧရိယာကို ကလစ်နှိပ်လို့လည်း ရပါတယ်။

၅။ Area drop-down list ထဲကနေ Data ကို ရွေးချယ်ပါ။

၆။ Background drop-down palette ထဲကနေ အရောင်အသစ်တစ်ခုကို ရွေးချယ်ပါ။

၇။ အဲဒီနောက် text color အသစ်တစ်ခုကို ရွေးချယ်ပါ။ Sample window မှာ style အသစ်ကို ပြသပေးပါလိမ့်မယ်။

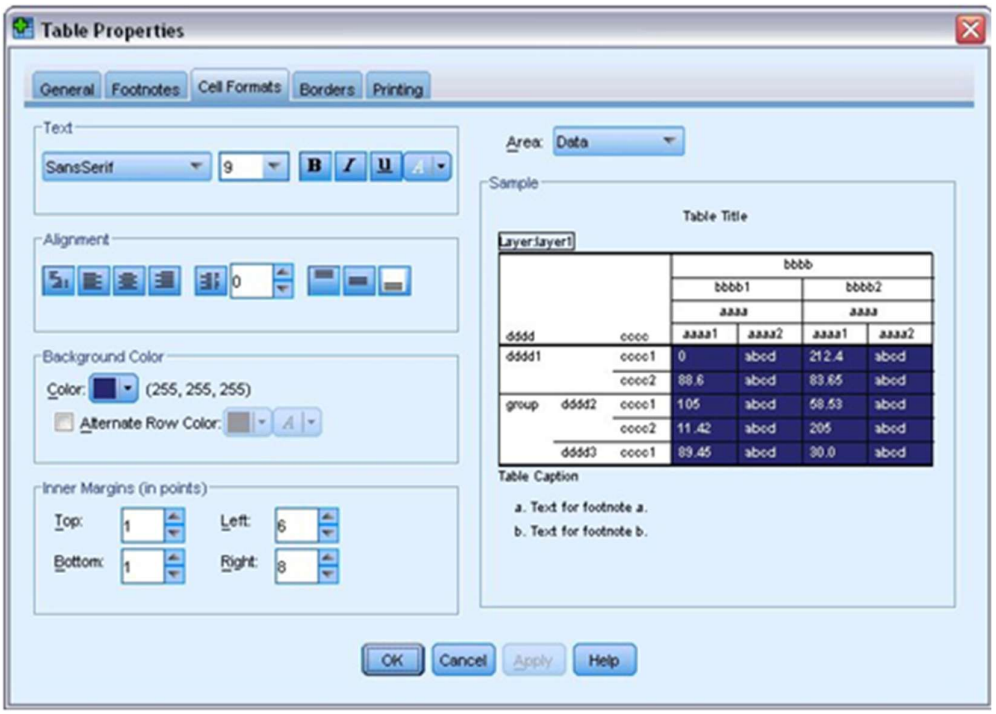


Figure 49. Changing table cell formats

၈။ TableLooks dialog box ဆီ ပြန်သွားဖို့ OK ကို နှိပ်ပါ။ သင့်ရဲ့ style အသစ်ကို save လုပ်ထားနိုင်ပြီး၊ အဲဒီလိုလုပ်ထားခြင်းဖြင့် နောင်လုပ်မယ့် tables တွေမှာ အလွယ်တကူ ပြန်သုံးနိုင်မှာပါ။

၉။ Save As ကို နှိပ်ပါ။

၁၀။ သိမ်းချင်တဲ့ target directory ဆီသွားပြီး၊ File Name text box ထဲမှာ သင့် style အသစ်အတွက် နာမည်တစ်ခု ပေးပါ။

၁၁။ Save ကို နှိပ်ပါ။

၁၂။ ပြောင်းလဲမှုတွေကို အသက်ဝင်စေပြီး Viewer ဆီ ပြန်သွားဖို့ OK ကို နှိပ်ပါ။ အခုဆိုရင် table ထဲမှာ သင်သတ်မှတ်ထားတဲ့ custom formatting တွေ ပါဝင်နေပြီ ဖြစ်ပါတယ်။

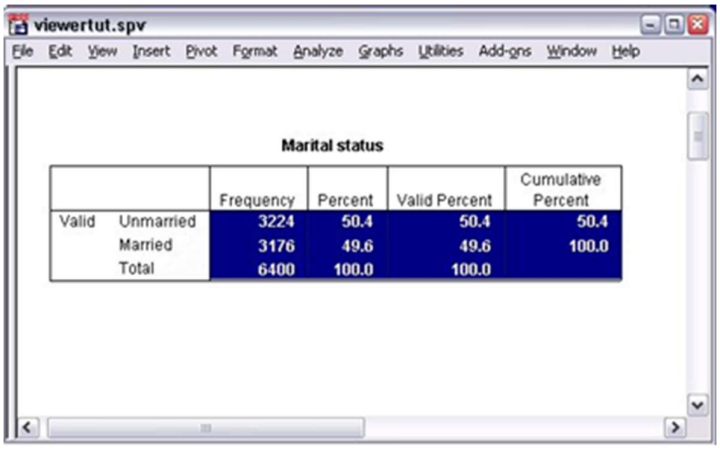

viewertut.spv
File Edit View Insert Pivot Format Analyze Graphs Utilities Add-ons Window Help
Marital status
Frequency Percent Valid Percent Cumulative Percent
Valid Unmarried 3224 50.4 50.4 50.4
Married 3176 49.6 49.6 100.0
Total 6400 100.0 100.0

Figure 50. Custom TableLook

Changing the Default Table Formats

Table တစ်ခုကို ဖန်တီးပြီးသွားမှ format ကို ပြောင်းလဲလို့ ရတယ်ဆိုပေမယ့်၊ table အသစ်တစ်ခု ဖန်တီးတိုင်း အမြဲတမ်း format ကို လိုက်ပြောင်းနေရမယ့်အစား default TableLook ကို ပြောင်းလဲထားလိုက်တာက ပိုပြီး အဆင်ပြေ (efficient) ပါလိမ့်မယ်။ Pivot tables တွေအတွက် default TableLook style ကို ပြောင်းလဲဖို့၊ menus တွေထဲကနေ အောက်ပါအတိုင်း ရွေးချယ်ပါ- Edit > Options...

၁။ Options dialog box ထဲက Pivot Tables tab ကို နှိပ်ပါ။

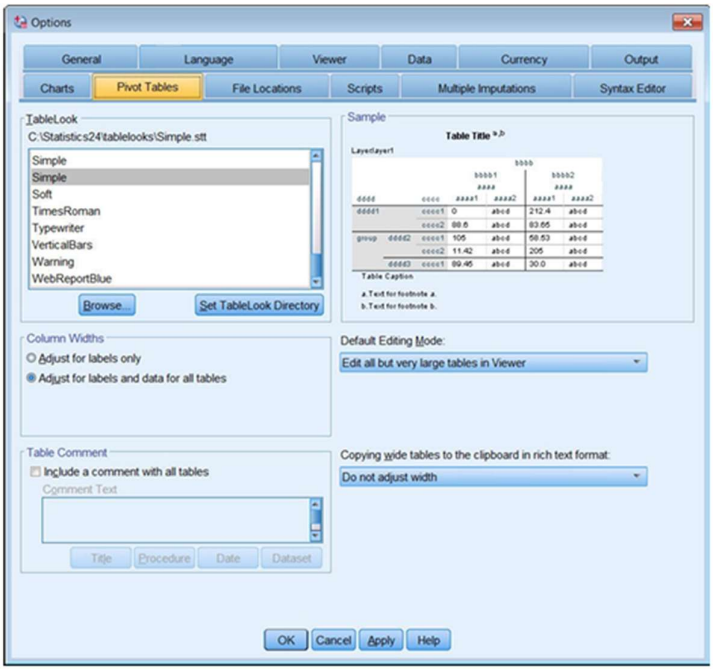


Figure 51. Options dialog box

၂။ Tables အသစ်တွေအားလုံးအတွက် သင်အသုံးပြုချင်တဲ့ TableLook style ကို ရွေးချယ်ပါ။ ညာဘက်က Sample window မှာ TableLook တစ်ခုချင်းစီရဲ့ preview ကို ပြသပေးပါတယ်။

၃။ Settings တွေကို save လုပ်ပြီး dialog box ကို ပိတ်ဖို့ OK ကို နှိပ်ပါ။

Default TableLook ကို ပြောင်းလဲပြီးနောက်ပိုင်း သင်ဖန်တီးသမျှ tables အားလုံးဟာ formatting အသစ်အတိုင်း အလိုအလျောက် ဖြစ်နေမှာပါ။

Customizing the Initial Display Settings

Initial display settings တွေထဲမှာ Viewer ထဲက objects တွေရဲ့ alignment တွေ၊ objects တွေကို default အနေနဲ့ ပြ (shown) မလား ဖျောက် (hidden) မလားဆိုတာနဲ့၊ Viewer window ရဲ့ အကျယ် (width) တို့ ပါဝင်ပါတယ်။

ဒီ settings တွေကို ပြောင်းလဲဖို့အတွက်-

၁။ menus တွေထဲကနေ Edit > Options... ကို ရွေးချယ်ပါ။

၂။ Viewer tab ကို နှိပ်ပါ။

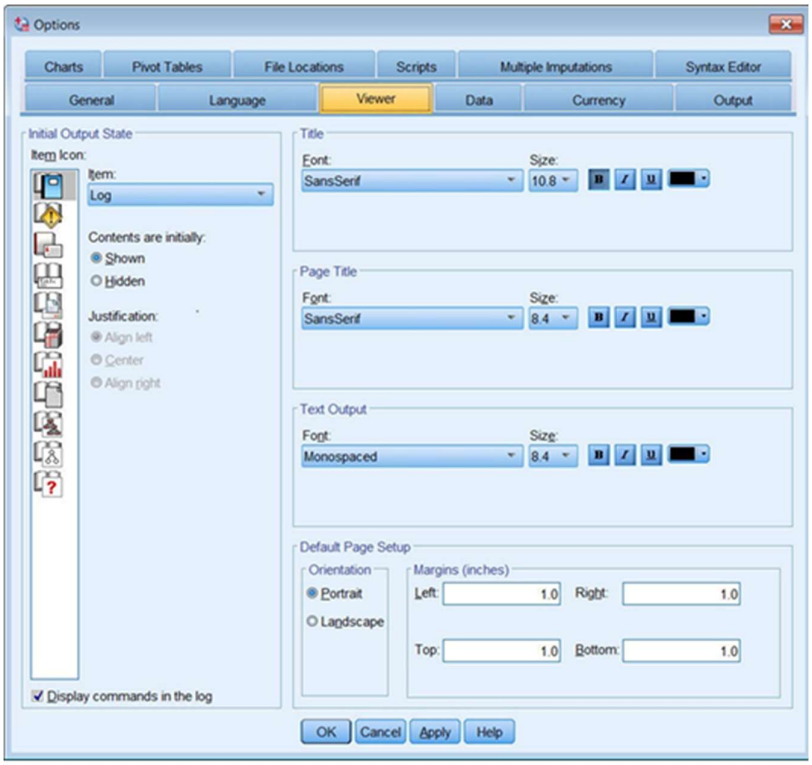


Figure 52. Viewer options

ဒီ settings တွေက object တစ်ခုချင်းစီအပေါ်မှာ သက်ရောက်မှုရှိတာပါ။ ဥပမာအားဖြင့်၊ သင်က tables တွေ ပြသတဲ့ပုံစံကို မပြောင်းလဲဘဲ charts တွေ ပြသတဲ့ပုံစံကို သီးသန့် customize လုပ်လို့ ရပါတယ်။ သင် customize လုပ်ချင်တဲ့ object ကို ရွေးချယ်ပြီး ပြောင်းလဲမှုတွေ ပြုလုပ်ရုံပါပဲ။

၃။ Title ရဲ့ settings တွေကို ကြည့်ဖို့ Title icon ကို နှိပ်ပါ။

၄။ Titles အားလုံးကို Viewer ရဲ့ အလယ် (horizontal center) မှာ ပြသဖို့ Center ကို နှိပ်ပါ။ သင့်ရဲ့ output ကို ရှုပ်ထွေးစေတတ်တဲ့ log တွေနဲ့ warning messages တွေလို elements တွေကိုလည်း ဖျောက်ထားနိုင်ပါတယ်။ Icon တစ်ခုကို double-click နှိပ်လိုက်တာနဲ့ အဲဒီ object ရဲ့ display property ကို အလိုအလျောက် ပြောင်းလဲသွားစေပါတယ်။

၅။ Output ထဲမှာ warning messages တွေကို ဖျောက်ဖို့ Warnings icon ကို double-click နှိပ်ပါ။

၆။ ပြောင်းလဲမှုတွေကို save လုပ်ပြီး dialog box ကို ပိတ်ဖို့ OK ကို နှိပ်ပါ။

Displaying Variable and Value Labels

အခြေအနေအများစုမှာ variables တွေနဲ့ values တွေအတွက် labels တွေကို ပြသတာက variable name တွေ ဒါမှမဟုတ် အမှန်တကယ် data value တွေကို ပြသတာထက် ပိုပြီး ထိရောက်မှုရှိပါတယ်။ ဒါပေမယ့် နာမည်တွေ (names) ရော labels တွေရော နှစ်မျိုးလုံးကို ပြသချင်တဲ့ အခြေအနေမျိုးလည်း ရှိနိုင်ပါတယ်။

၁။ menus တွေထဲကနေ Edit > Options... ကို ရွေးချယ်ပါ။

၂။ Output Labels tab ကို နှိပ်ပါ။

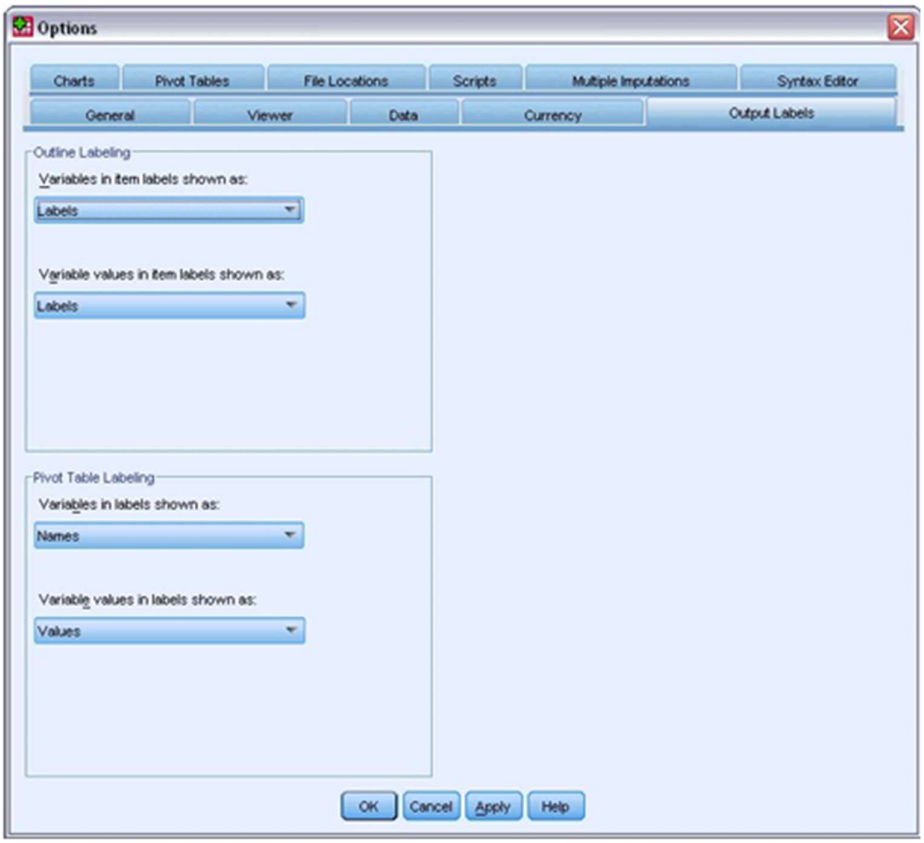


Figure 53. Pivot Table Labeling settings

Outline နဲ့ contents panes တွေအတွက် မတူညီတဲ့ settings တွေကို သတ်မှတ်ပေးနိုင်ပါတယ်။ ဥပမာအနေနဲ့၊ outline ထဲမှာ labels တွေကို ပြသပြီး၊ contents ထဲမှာတော့ variable names တွေနဲ့ data values တွေကို ပြသချင်တယ်ဆိုရင်-

၃။ Pivot Table Labeling group ထဲက၊ Variables in Labels drop-down list မှာ Labels အစား variable names တွေကို ပြသဖို့ Names ကို ရွေးချယ်ပါ။

၄။ ပြီးရင်၊ Variable Values in Labels drop-down list မှာ labels တွေအစား data values တွေကို ပြသဖို့ Values ကို ရွေးချယ်ပါ။

ဒီ session ထဲမှာ နောက်ထပ်ထုတ်မယ့် tables တွေမှာ အခုပြောင်းလဲမှုတွေကို မြင်တွေ့ရမှာ ဖြစ်ပါတယ်။

marital

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3224	50.4	50.4	50.4
	1	3176	49.6	49.6	100.0
	Total	6400	100.0	100.0	

Figure 54. Variable names and values displayed

Using Results in Other Applications

သင့်ရဲ့ results တွေကို တခြား applications တွေမှာလည်း အသုံးပြုနိုင်ပါတယ်။ ဥပမာအနေနဲ့ table ဒါမှမဟုတ် chart တစ်ခုကို presentation ဒါမှမဟုတ် report တစ်ခုထဲမှာ ထည့်သွင်းချင်တာမျိုးပေါ့။

အောက်မှာပြထားတဲ့ ဥပမာတွေက Microsoft Word အတွက် သီးသန့်ဖြစ်ပေမယ့်၊ တခြား word processing applications တွေမှာလည်း အလားတူ အလုပ်လုပ်နိုင်ပါတယ်။

Pasting Results as Word Tables

Pivot tables တွေကို Word ထဲကို မူလ Word tables တွေအနေနဲ့ paste လုပ်ထည့်နိုင်ပါတယ်။ Font sizes တွေနဲ့ အရောင်တွေလိုမျိုး table attributes အားလုံးကို ထိန်းသိမ်းထားပေးပါတယ်။ Table ကို Word table format နဲ့ paste လုပ်လိုက်တာဖြစ်တဲ့အတွက်၊ Word ထဲက တခြား table တွေလိုပဲ သူ့ကိုလည်း edit လုပ်နိုင်ပါတယ်။

၁။ Viewer ထဲက table တစ်ခုကို ရွေးချယ်ဖို့ ကလစ်နှိပ်ပါ။

၂။ menus တွေထဲကနေ Edit > Copy ကို ရွေးချယ်ပါ။

၃။ သင့်ရဲ့ word processing application ကို ဖွင့်ပါ။

၄။ Word processor ရဲ့ menus တွေထဲကနေ Edit > Paste Special... ကို ရွေးချယ်ပါ။

၅။ Paste Special dialog box ထဲမှာ Formatted Text (RTF) ကို ရွေးချယ်ပါ။

၆။ သင့် document ထဲကို results တွေ paste လုပ်ဖို့ OK ကို နှိပ်ပါ။

အခုဆိုရင် table ကို သင့် document ထဲမှာ ပြသနေပြီ ဖြစ်ပါတယ်။ Custom formatting တွေ လုပ်တာ၊ data တွေကို edit လုပ်တာနဲ့၊ သင့်လိုအပ်ချက်နဲ့ ကိုက်ညီအောင် table အရွယ်အစား ပြောင်းလဲတာတွေကို လုပ်ဆောင်နိုင်ပါတယ်။

Pasting Results as Text

Pivot tables တွေကို တခြား applications တွေဆီ plain text အနေနဲ့လည်း copy ကူးထည့်နိုင်ပါတယ်။ ဒီနည်းလမ်းမှာတော့ formatting styles တွေကို ထိန်းသိမ်းထားမှာ မဟုတ်ပေမယ့်၊ target application ထဲကို paste လုပ်ပြီးတဲ့အခါမှာ table data တွေကို edit လုပ်နိုင်ပါတယ်။

၁။ Viewer ထဲက table တစ်ခုကို ရွေးချယ်ဖို့ ကလစ်နှိပ်ပါ။

၂။ menus တွေထဲကနေ Edit > Copy ကို ရွေးချယ်ပါ။

၃။ သင့်ရဲ့ word processing application ကို ဖွင့်ပါ။

၄။ Word processor ရဲ့ menus တွေထဲကနေ Edit > Paste Special... ကို ရွေးချယ်ပါ။

၅။ Paste Special dialog box ထဲမှာ Unformatted Text ကို ရွေးချယ်ပါ။

၆။ သင့် document ထဲကို results တွေ paste လုပ်ဖို့ OK ကို နှိပ်ပါ။

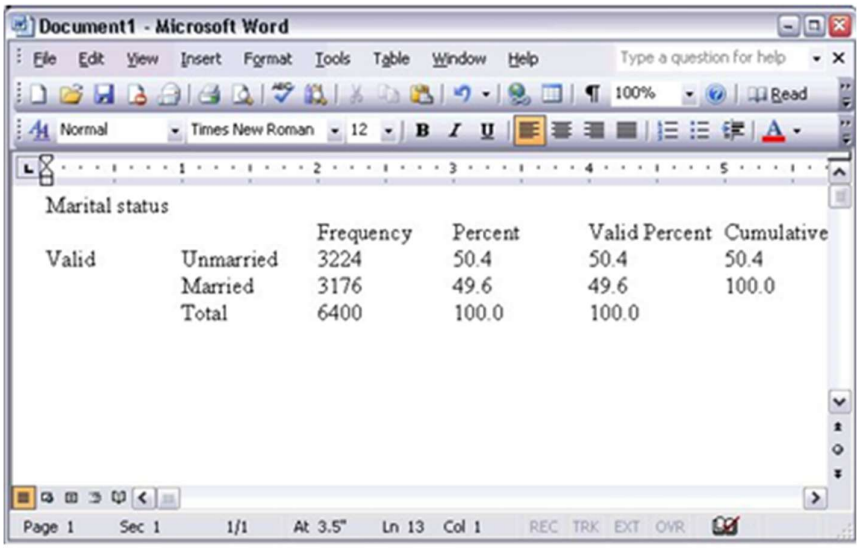

Document1 - Microsoft Word
File Edit View Insert Format Tools Table Window Help
Type a question for help
Normal Times New Roman 12
Marital status
Frequency Percent Valid Percent Cumulative
Valid Unmarried 3224 50.4 50.4 50.4
Married 3176 49.6 49.6 100.0
Total 6400 100.0 100.0
Page 1 Sec 1 1/1 At 3.5" Ln 13 Col 1

Figure 55. Pivot table displayed in Word

Table ရဲ့ column တစ်ခုစီကို tabs တွေနဲ့ ခွဲခြားထားပါတယ်။ သင့် word processing application ထဲက tab stops တွေကို ချိန်ညှိပြီး column အကျယ်တွေကို ပြောင်းလဲနိုင်ပါတယ်။

Exporting Results to Microsoft Word, PowerPoint, and Excel Files

Results တွေကို Microsoft Word, PowerPoint ဒါမှမဟုတ် Excel file အဖြစ် export လုပ်နိုင်ပါတယ်။ ရွေးချယ်ထားတဲ့ items တွေကိုဖြစ်စေ၊ Viewer ထဲက items အားလုံးကိုဖြစ်စေ export လုပ်နိုင်ပါတယ်။ ဒီအပိုင်းမှာတော့ msouttut.spv နဲ့ demo.sav ဆိုတဲ့ files တွေကို သုံးသွားမှာပါ။ ပိုပြီးသိချင်တယ်ဆိုရင် Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ ကြည့်နိုင်ပါတယ်။

မှတ်ချက် - PowerPoint သို့ export လုပ်ခြင်းကို Windows operating systems တွေမှာသာ အသုံးပြုနိုင်ပြီး Student Version မှာတော့ မရနိုင်ပါဘူး။ Viewer ရဲ့ outline pane ထဲမှာ၊ သင်

export လုပ်ချင်တဲ့ သီးသန့် items တွေကို ရွေးချယ်နိုင်သလို၊ items အားလုံးကိုဖြစ်စေ၊ မြင်နေရတဲ့ (visible) items အားလုံးကိုဖြစ်စေ export လုပ်နိုင်ပါတယ်။

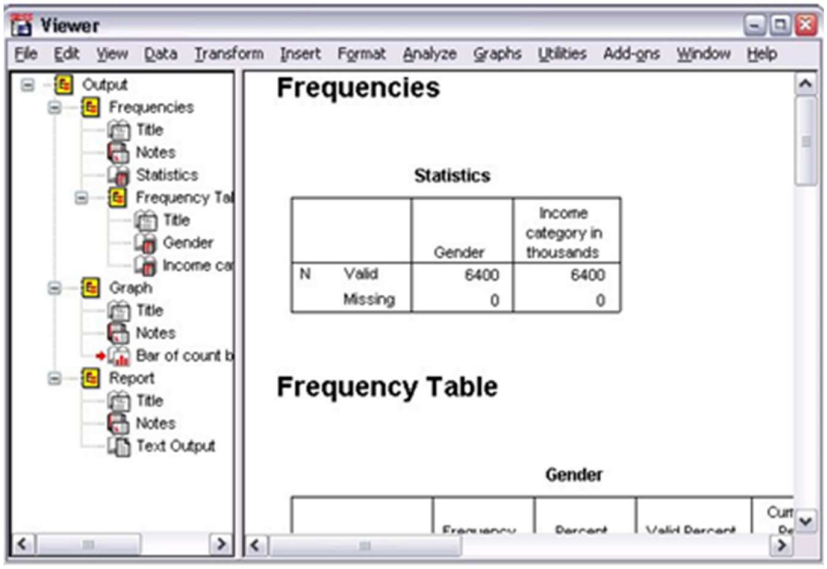


Figure 56. Viewer

၁။ Viewer ရဲ့ menus တွေထဲကနေ File > Export... ကို ရွေးချယ်ပါ။

Viewer ထဲက objects အားလုံးကို export လုပ်မယ့်အစား၊ မြင်နေရတဲ့ objects တွေ (outline pane ထဲက ဖွင့်ထားတဲ့ စာအုပ်ပုံလေးတွေ) ကိုသာဖြစ်စေ၊ ဒါမှမဟုတ် outline pane ထဲမှာ သင်ရွေးချယ်ထားတဲ့ အရာတွေကိုသာဖြစ်စေ export လုပ်ဖို့ ရွေးချယ်နိုင်ပါတယ်။ တကယ်လို့ သင်က outline pane ထဲမှာ ဘာ items ကိုမှ မရွေးချယ်ထားဘူးဆိုရင်တော့ ရွေးချယ်ထားတဲ့ objects တွေကို export လုပ်ဖို့ဆိုတဲ့ option က ရွေးလို့ရမှာ မဟုတ်ပါဘူး။

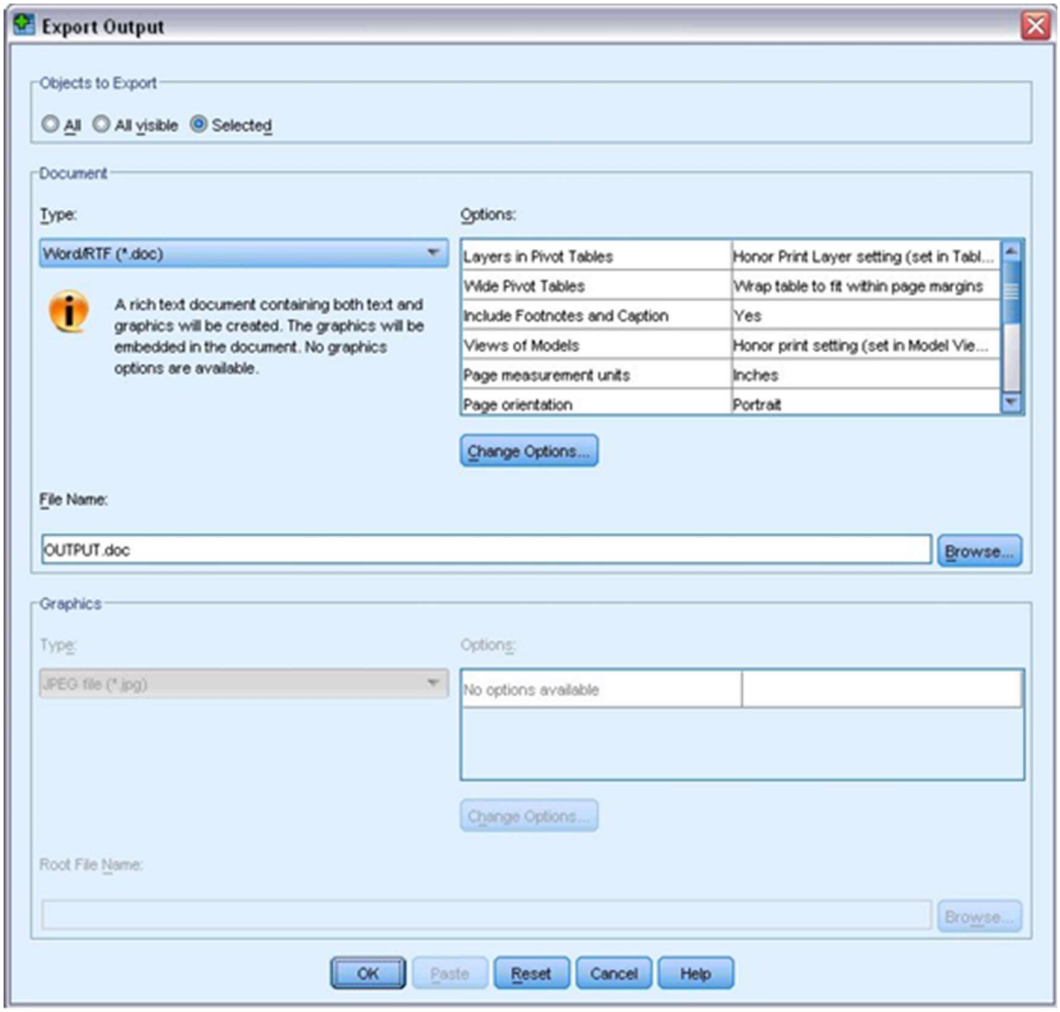


Figure 57. Export Output dialog box

၂။ Objects to Export group ထဲမှာ All ကို ရွေးချယ်ပါ။

၃။ Type drop-down list ထဲကနေ Word/RTF file (*.doc) ကို ရွေးချယ်ပါ။

၄။ Word file ကို ထုတ်ယူ (generate) ဖို့ OK ကို နှိပ်ပါ။ ထွက်လာတဲ့ဖိုင်ကို Word ထဲမှာ ဖွင့်လိုက်တဲ့အခါ၊ results တွေကို ဘယ်လို export လုပ်ထားတယ်ဆိုတာ မြင်ရမှာပါ။

မြင်ရတဲ့ objects တွေမဟုတ်တဲ့ Notes တွေက သင်က objects အားလုံးကို export လုပ်ဖို့ ရွေးချယ်ခဲ့တဲ့အတွက် Word ထဲမှာ ပါလာတာ ဖြစ်ပါတယ်။ Pivot tables တွေက Word tables တွေအဖြစ် ပြောင်းသွားပြီး၊ fonts တွေ၊ အရောင်တွေ၊ borders တွေ စသဖြင့် မူရင်း pivot table ရဲ့ formatting အားလုံးကို အပြည့်အဝ ထိန်းသိမ်းထားပါတယ်။

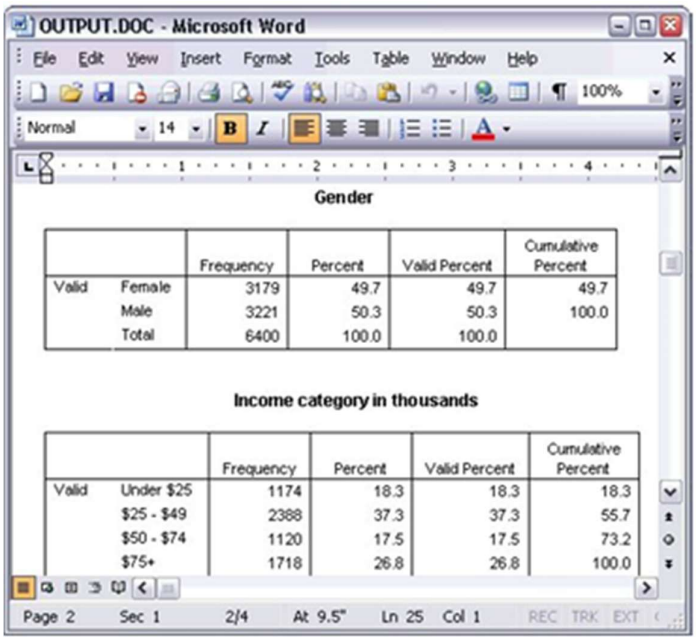

OUTPUT.DOC - Microsoft Word
File Edit View Insert Format Tools Table Window Help
Normal 14
Gender
Frequency Percent Valid Percent Cumulative Percent
Valid Female 3179 49.7 49.7 49.7
Male 3221 50.3 50.3 100.0
Total 6400 100.0 100.0
Income category in thousands
Frequency Percent Valid Percent Cumulative Percent
Valid Under $25 1174 18.3 18.3 18.3
$25 - $49 2388 37.3 37.3 55.7
$50 - $74 1120 17.5 17.5 73.2
$75+ 1718 26.8 26.8 100.0
Page 2 Sec 1 2/4 At 9.5" Ln 25 Col 1

Figure 58. Pivot tables in Word

Charts တွေကိုတော့ Word document ထဲမှာ graphic images တွေအနေနဲ့ ထည့်သွင်းပေးပါတယ်။

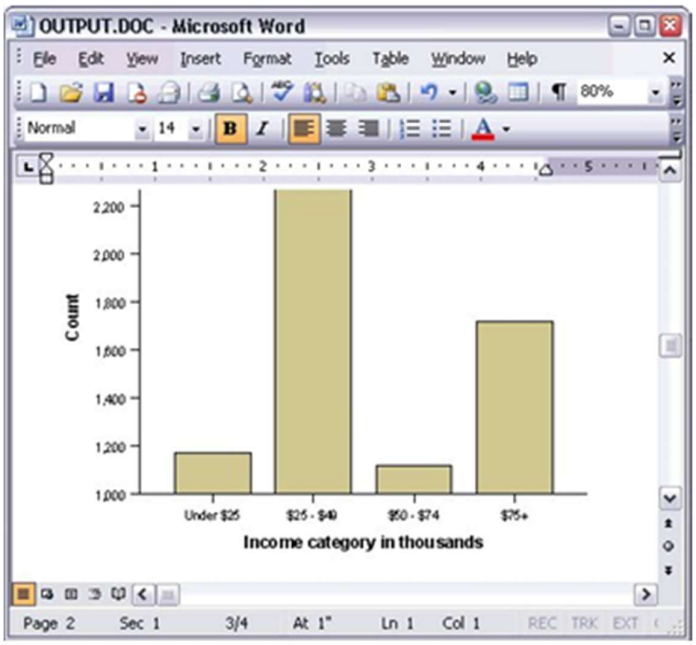


Figure 59. Charts in Word

Text output ကိုတော့ Viewer ထဲက text object မှာ သုံးခဲ့တဲ့ font အတိုင်း ပြသပေးပါတယ်။ နေရာချထားမှု (alignment) မှန်ကန်စေဖို့အတွက်၊ text output မှာ fixed-pitch (monospaced) font ကို အသုံးပြုသင့်ပါတယ်။

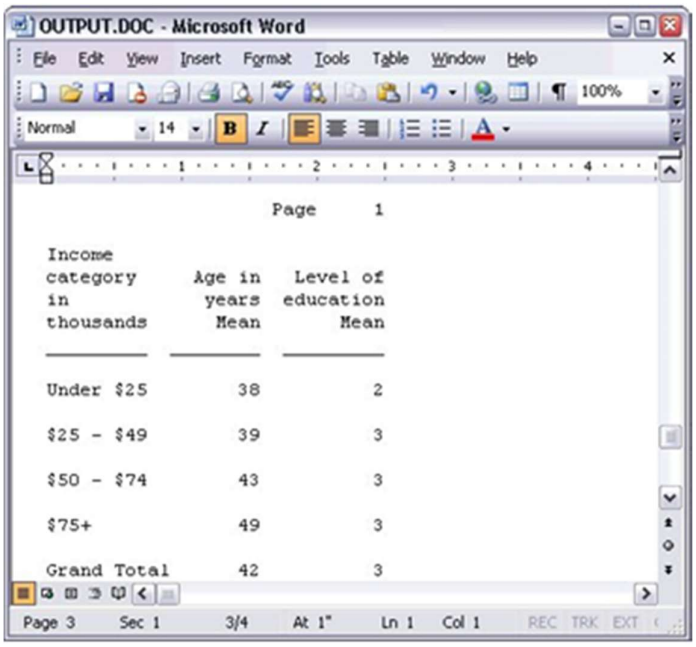

OUTPUT.DOC - Microsoft Word
File Edit View Insert Format Tools Table Window Help
Page 1
Income category in thousands
Age in years Mean
Level of education Mean
Under $25 38 2
$25 - $49 39 3
$50 - $74 43 3
$75+ 49 3
Grand Total 42 3
Page 3 Sec 1 3/4 At 1" Ln 1 Col 1

Figure 60. Text output in Word

တကယ်လို့ သင်က PowerPoint file အနေနဲ့ export လုပ်မယ်ဆိုရင်တော့၊ export လုပ်လိုက်တဲ့ item တစ်ခုစီကို slide တစ်ခုစီမှာ ထည့်သွင်းပေးမှာပါ။ PowerPoint ဆီ export လုပ်လိုက်တဲ့ Pivot tables တွေဟာ Word tables တွေလိုပဲ မူလ pivot table ရဲ့ formatting တွေဖြစ်တဲ့ fonts တွေ၊ colors တွေ၊ borders တွေစတာတွေကို အားလုံး ထိန်းသိမ်းထားပါတယ်။

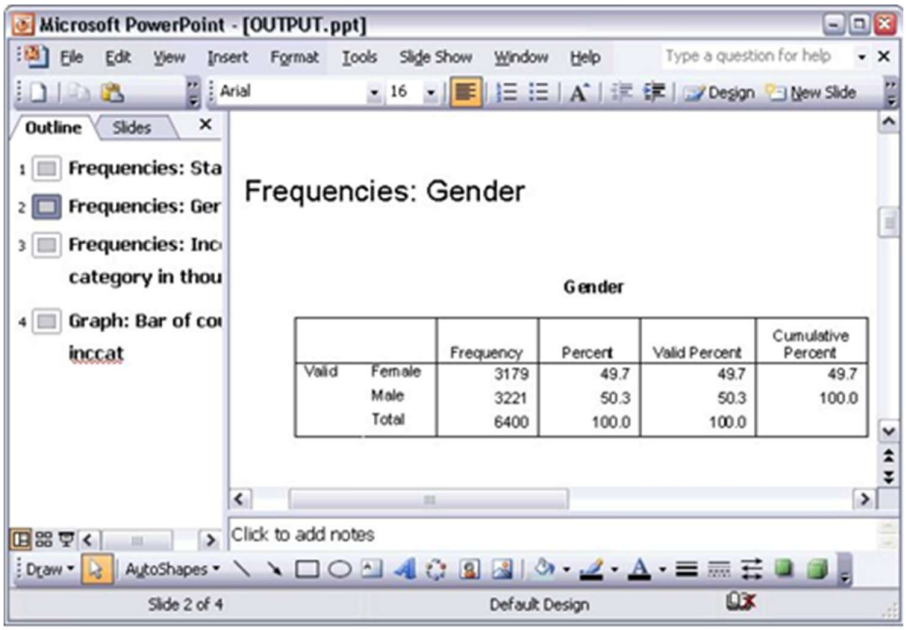


Figure 61. Pivot tables in PowerPoint

PowerPoint ဆီ export လုပ်ဖို့ ရွေးချယ်ထားတဲ့ Charts တွေကိုတော့ PowerPoint file ထဲမှာ embedded အနေနဲ့ ထည့်သွင်းပေးပါတယ်။

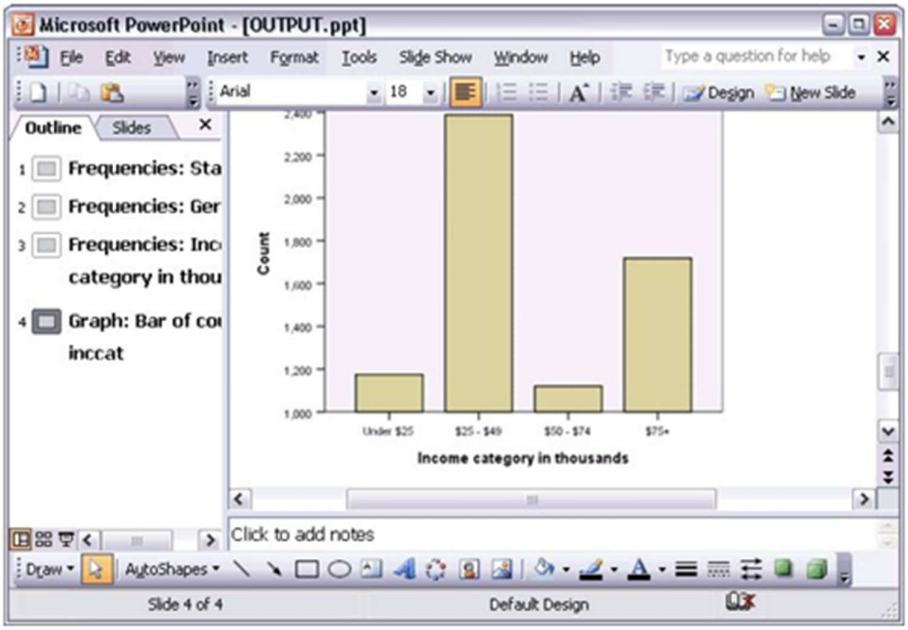


Figure 62. Charts in PowerPoint

မှတ်ချက် - PowerPoint သို့ export လုပ်ခြင်းကို Windows operating systems တွေမှာသာ ရရှိနိုင်ပြီး Student Version မှာ အသုံးမပြုနိုင်ပါဘူး။ တကယ်လို့ သင်က Excel file အနေနဲ့ export လုပ်မယ်ဆိုရင်တော့ results တွေကို မတူညီတဲ့ပုံစံနဲ့ export လုပ်ပေးမှာပါ။

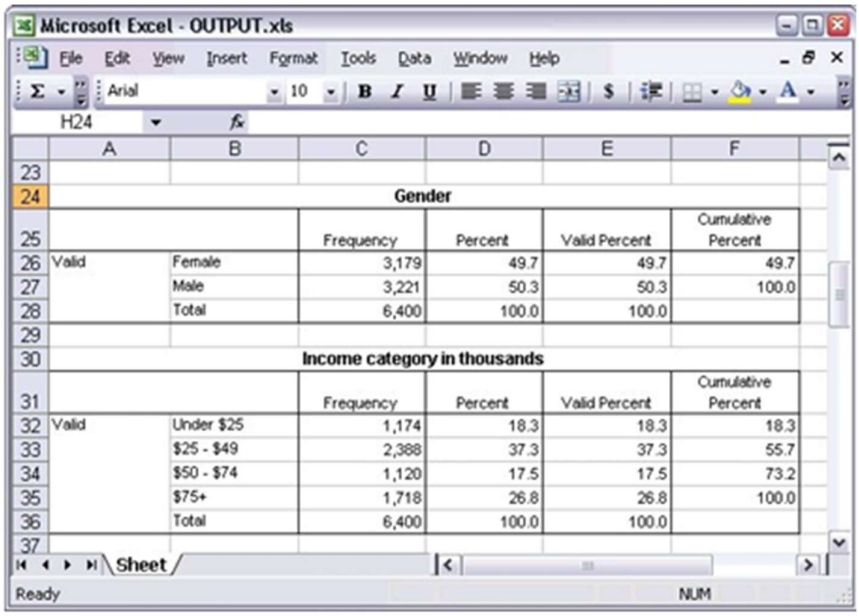
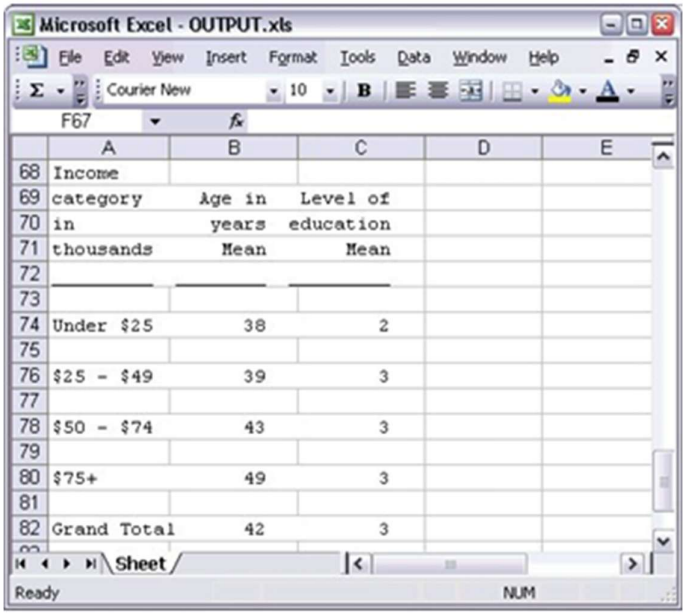

Microsoft Excel - OUTPUT.xls
File Edit View Insert Format Tools Data Window Help
H24
Gender
Frequency Percent Valid Percent Cumulative Percent
Valid Female 3,179 49.7 49.7 49.7
Male 3,221 50.3 50.3 100.0
Total 6,400 100.0 100.0
Income category in thousands
Frequency Percent Valid Percent Cumulative Percent
Valid Under $25 1,174 18.3 18.3 18.3
$25 - $49 2,388 37.3 37.3 55.7
$50 - $74 1,120 17.5 17.5 73.2
$75+ 1,718 26.8 26.8 100.0
Total 6,400 100.0 100.0
Sheet
Ready
NUM

Figure 63. Output.xls in Excel

Pivot table ရဲ့ rows တွေ၊ columns တွေနဲ့ cells တွေက Excel ရဲ့ rows တွေ၊ columns တွေနဲ့ cells တွေအဖြစ် ပြောင်းလဲသွားမှာပါ။ Text output ထဲက စာကြောင်းတစ်ကြောင်းစီက Excel file ထဲမှာ row တစ်ခု ဖြစ်သွားပြီး၊ အဲဒီစာကြောင်းရဲ့ အကြောင်းအရာတစ်ခုလုံးကို cell တစ်ခုတည်းမှာပဲ ထည့်သွင်းပေးမှာ ဖြစ်ပါတယ်။



	A	B	C	D	E
68	Income				
69	category	Age in	Level of		
70	in	years	education		
71	thousands	Mean	Mean		
72	_____	_____	_____		
73					
74	Under $25	38	2		
75					
76	$25 - $49	39	3		
77					
78	$50 - $74	43	3		
79					
80	$75+	49	3		
81					
82	Grand Total	42	3		

Figure 64. Text output in Excel

Exporting Results to PDF

Viewer ထဲက items အားလုံးကိုဖြစ်စေ၊ ရွေးချယ်ထားတဲ့ items တွေကိုဖြစ်စေ PDF (portable document format) file အဖြစ် export လုပ်နိုင်ပါတယ်။

၁။ သင် PDF အဖြစ် export လုပ်ချင်တဲ့ result ရှိနေတဲ့ Viewer window ရဲ့ menus တွေထဲကနေ File > Export... ကို ရွေးချယ်ပါ။

၂။ Export Output dialog box ထဲက Export Format File Type drop-down list မှာ Portable Document Format ကို ရွေးချယ်ပါ။

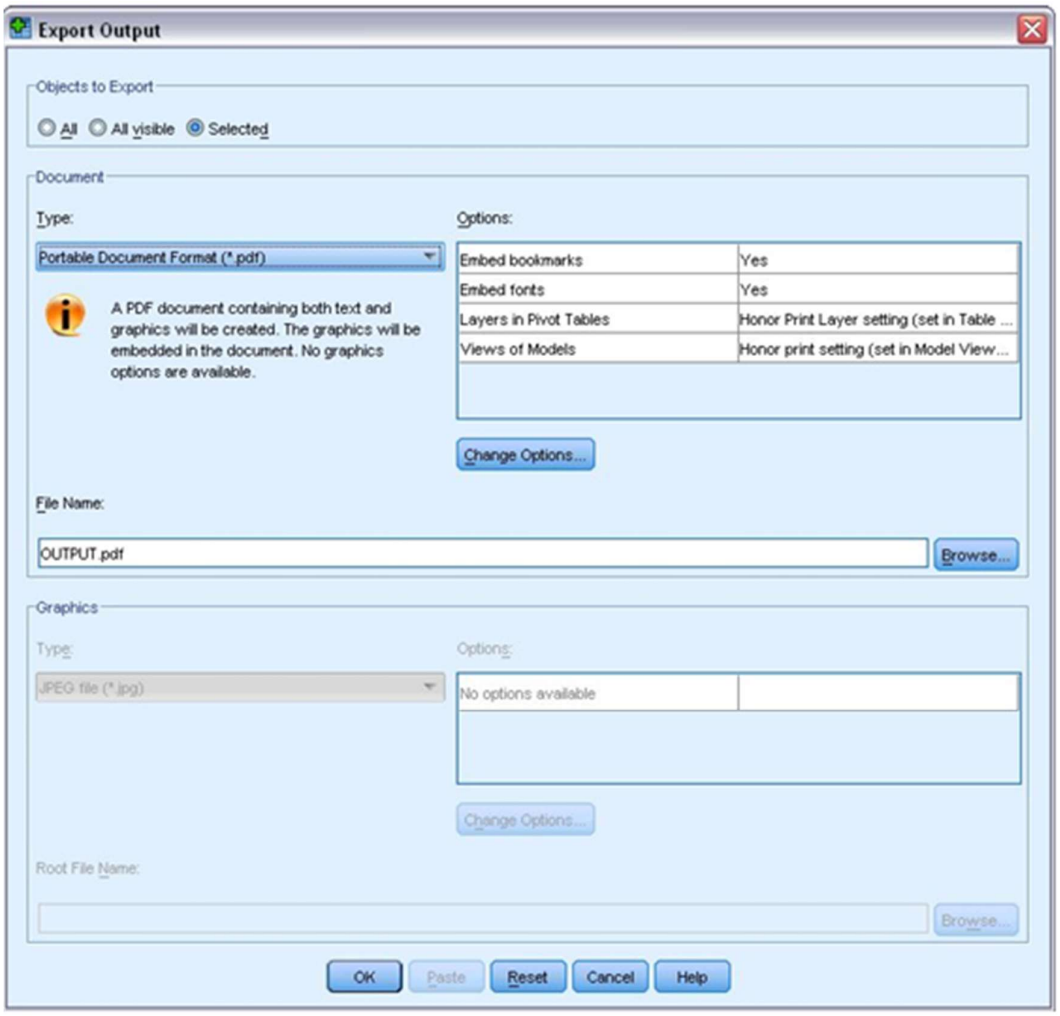


Figure 65. Export Output dialog box

- လွယ်ကူစွာ သွားလာရှာဖွေနိုင်ဖို့အတွက် Viewer document ရဲ့ outline pane ကို PDF file ထဲမှာ bookmarks တွေအနေနဲ့ ပြောင်းလဲပေးပါတယ်။
- PDF documents တွေမှာပါတဲ့ page size, orientation, margins, page headers နဲ့ footers ရဲ့ အကြောင်းအရာနဲ့ ပြသမှု၊ အပြင် print ထုတ်မယ့် chart size တွေကို page setup options (Viewer window ထဲက File menu > Page Setup) ကနေ ထိန်းချုပ်နိုင်ပါတယ်။
- PDF document ရဲ့ resolution (DPI) က default ဒါမှမဟုတ် လက်ရှိရွေးချယ်ထားတဲ့ printer ရဲ့ လက်ရှိ resolution setting အတိုင်းဖြစ်ပါတယ် (Page Setup ကို သုံးပြီး ပြောင်းလဲနိုင်ပါတယ်) ။ အများဆုံး resolution ကတော့ 1200 DPI ပါ။ တကယ်လို့ printer setting က အဲဒီထက် ပိုမြင့်နေမယ်ဆိုရင်တောင် PDF document ရဲ့ resolution က 1200 DPI ပဲ ဖြစ်နေမှာပါ။ မှတ်ချက် - High-resolution documents တွေကို resolution နိမ့်တဲ့ printers တွေနဲ့ print ထုတ်တဲ့အခါ ရလဒ် သိပ်မကောင်းတာမျိုး ဖြစ်နိုင်ပါတယ်။

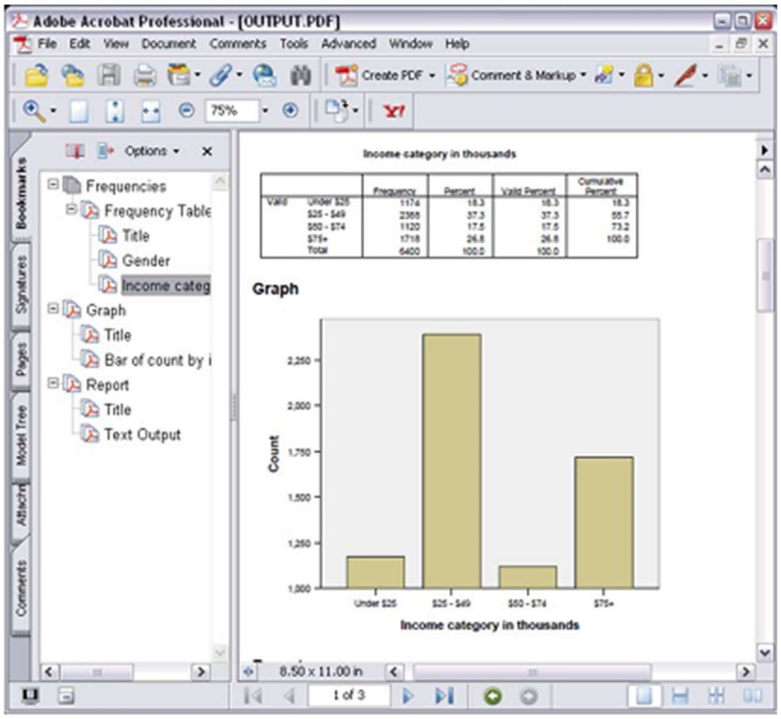


Figure 66. PDF file with bookmarks

Exporting Results to HTML

Results တွေကို HTML (hypertext markup language) အဖြစ်လည်း export လုပ်နိုင်ပါသေးတယ်။ HTML အနေနဲ့ save လုပ်တဲ့အခါ၊ graphic မဟုတ်တဲ့ output အားလုံးကို HTML file တစ်ခုတည်းထဲမှာပေါင်းပြီး export လုပ်ပေးမှာပါ။

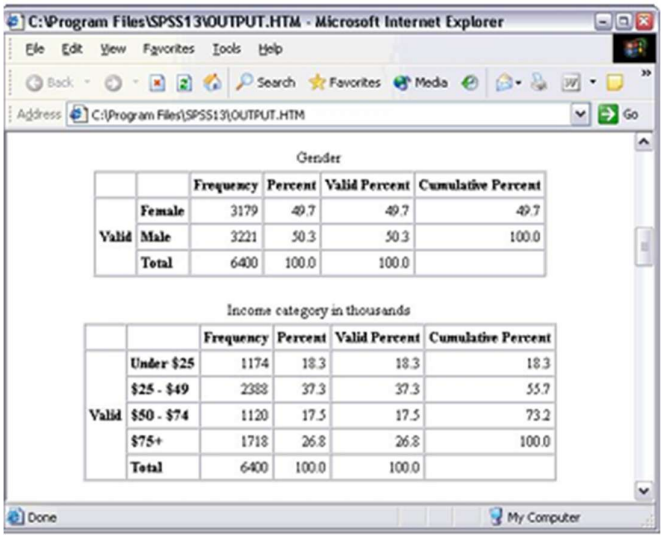

C:\Program Files\SPSS13\OUTPUT.HTM - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address C:\Program Files\SPSS13\OUTPUT.HTM
Gender
Frequency Percent Valid Percent Cumulative Percent
Female 3179 49.7 49.7 49.7
Valid Male 3221 50.3 50.3 100.0
Total 6400 100.0 100.0
Income category in thousands
Frequency Percent Valid Percent Cumulative Percent
Under $25 1174 18.3 18.3 18.3
$25 - $49 2388 37.3 37.3 55.7
Valid $50 - $74 1120 17.5 17.5 73.2
$75+ 1718 26.8 26.8 100.0
Total 6400 100.0 100.0
Done
My Computer

Figure 67. Output.htm in Web browser

HTML အဖြစ် export လုပ်တဲ့အခါ charts တွေကိုလည်း export လုပ်နိုင်ပေမယ့် ဖိုင်တစ်ခုတည်းထဲမှာတော့ မဟုတ်ပါဘူး။

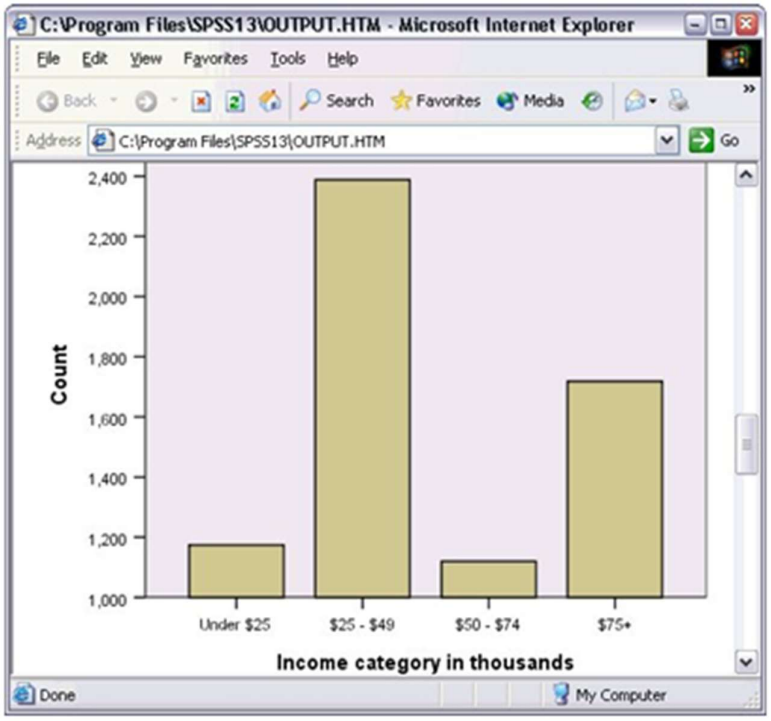


Figure 68. Chart in HTML

Chart တစ်ခုစီကို သင်သတ်မှတ်ထားတဲ့ format နဲ့ ဖိုင်တစ်ခုအနေနဲ့ save လုပ်ပေးမှာဖြစ်ပြီး၊ ဒီ graphics files တွေရဲ့ လမ်းညွှန် (references) တွေကိုတော့ HTML ထဲမှာ ထည့်သွင်းပေးထားမှာပါ။ Charts အားလုံးကို (ဒါမှမဟုတ် ရွေးချယ်ထားတဲ့ charts တွေကို) သီးခြား graphics files တွေအဖြစ် export လုပ်နိုင်တဲ့ option လည်း ရှိပါတယ်။

Chapter 7. Working with Syntax

စွမ်းအားပြည့်ဝတဲ့ command language ကိုသုံးပြီး အများဆုံးလုပ်လေ့ရှိတဲ့ tasks တော်တော်များများကို save ထားတာ၊ automate လုပ်တာတွေ လုပ်နိုင်ပါတယ်။ သူက menus တွေနဲ့ dialog boxes တွေမှာ မပါဝင်တဲ့ functionality တချို့ကိုလည်း ထောက်ပံ့ပေးထားပါသေးတယ်။ Commands အများစုကိုတော့ menus တွေ၊ dialog boxes တွေကနေတစ်ဆင့် အသုံးပြုနိုင်ပါတယ်။ ဒါပေမယ့် တချို့ commands တွေနဲ့ options တွေကိုတော့ command language သုံးမှပဲ ရနိုင်မှာပါ။ Command language က သင့်ရဲ့ jobs တွေကို syntax file တစ်ခုအနေနဲ့ save လုပ်ခွင့်ပေးတဲ့အတွက် နောက်ပိုင်းမှာ သင့်ရဲ့ analysis ကို ထပ်ခါထပ်ခါ ပြန်လုပ်နိုင်မှာ ဖြစ်ပါတယ်။

Command syntax file ဆိုတာ ရိုးရှင်းစွာပြောရရင် IBM SPSS Statistics syntax commands တွေ ပါဝင်တဲ့ text file တစ်ခုပါပဲ။ သင်က syntax window တစ်ခုကိုဖွင့်ပြီး commands တွေကို တိုက်ရိုက် ရိုက်ထည့်နိုင်ပေမယ့်၊ အလုပ်တွေ တစ်စိတ်တစ်ပိုင်း ဒါမှမဟုတ် အားလုံးကို dialog boxes တွေကိုပဲ လုပ်ခိုင်းလိုက်တာက ပိုပြီး လွယ်ကူလေ့ရှိပါတယ်။

ဒီ chapter မှာပါတဲ့ ဥပမာတွေက demo.sav ဆိုတဲ့ data file ကို အသုံးပြုထားတာပါ။ ပိုပြီး သိချင်တယ်ဆိုရင်တော့ Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ သွားကြည့်နိုင်ပါတယ်။ မှတ်ချက် - Command syntax ကို Student Version မှာ အသုံးပြုနိုင်မှာ မဟုတ်ပါဘူး။

Pasting Syntax

Syntax တစ်ခုကို ဖန်တီးဖို့ အလွယ်ဆုံးနည်းလမ်းကတော့ dialog boxes အများစုမှာ ပါတဲ့ Paste button ကို အသုံးပြုခြင်းပဲ ဖြစ်ပါတယ်။

၁။ demo.sav ဆိုတဲ့ data file ကို ဖွင့်ပါ။ (ပိုသိချင်ရင် Chapter 10 က "Sample Files" မှာ ကြည့်ပါ)။

၂။ menus တွေထဲကနေ Analyze > Descriptive Statistics > Frequencies... ကို ရွေးချယ်ပါ။

၃။ Marital status [marital] ကို ရွေးချယ်ပြီး Variable(s) list ထဲကို ရွှေ့ထည့်ပါ။

၄။ Charts ကို နှိပ်ပါ။

၅။ Charts dialog box ထဲမှာ Bar charts ကို ရွေးချယ်ပါ။

၆။ Chart Values group ထဲမှာ Percentages ကို ရွေးချယ်ပါ။

၇။ Continue ကို နှိပ်ပါ။ Dialog box ထဲမှာ သင်ရွေးချယ်ခဲ့တဲ့အရာတွေကြောင့် ဖန်တီးလိုက်တဲ့ syntax ကို Syntax Editor ထဲ copy ကူးထည့်ဖို့ Paste ကို နှိပ်ပါ။

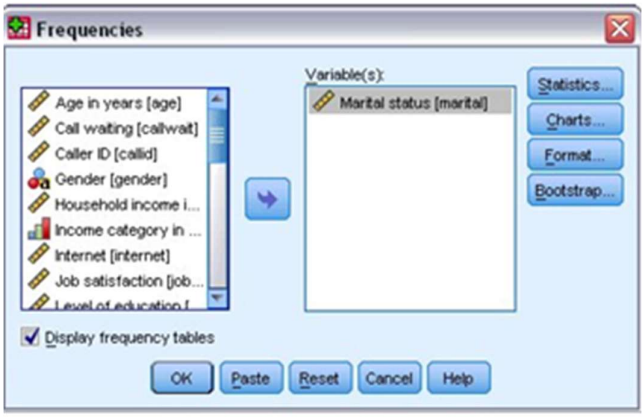


Figure 69. Frequencies dialog box

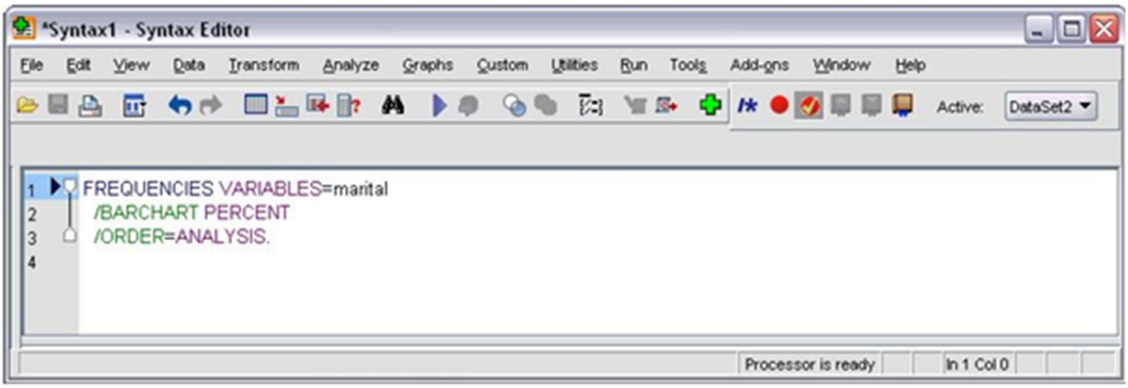

*Syntax1 - Syntax Editor
File Edit View Data Transform Analyze Graphs Custom Utilities Run Tools Add-ons Window Help
Active: DataSet2
1 FREQUENCIES VARIABLES=marital
2 /BARCHART PERCENT
3 /ORDER=ANALYSIS.
4
Processor is ready
ln 1 Col 0

Figure 70. Frequencies syntax

၈။ လက်ရှိ ပြသနေတဲ့ syntax ကို run ဖို့အတွက်၊ menus တွေထဲကနေ Run > Selection ကို ရွေးချယ်ပါ။

Editing Syntax

Syntax window ထဲမှာ၊ သင်က syntax ကို edit လုပ်နိုင်ပါတယ်။ ဥပမာအနေနဲ့၊ percentages တွေအစား frequencies တွေကို ပြသဖို့အတွက် / BARCHART ဆိုတဲ့ subcommand ကို ပြောင်းလဲနိုင်ပါတယ်။ (Subcommand တစ်ခုကို slash လေး (/) နဲ့ ညွှန်ပြလေ့ရှိပါတယ်)။ တကယ်လို့ သင်က frequencies တွေ ပြသဖို့အတွက် keyword ကို သိနေတယ်ဆိုရင် တိုက်ရိုက် ရိုက်ထည့်လိုက်လို့ ရပါတယ်။ မသိဘူးဆိုရင်တော့ cursor ကို subcommand name ရဲ့နောက် ဘယ်နေရာမှာမဆိုထားပြီး Ctrl+Spacebar ကို နှိပ်ခြင်းအားဖြင့် အဲဒီ subcommand အတွက်

အသုံးပြုနိုင်တဲ့ keywords စာရင်းကို ရယူနိုင်ပါတယ်။ ဒါက subcommand အတွက် auto-completion control ကို ပြသပေးမှာပါ။

BARCHART subcommand ထဲက PERCENT ဆိုတဲ့ keyword ကို ဖျက်လိုက်ပါ။ Ctrl-Spacebar ကို နှိပ်ပါ။

၁။ frequencies အတွက် FREQ လို့ နာမည်တပ်ထားတဲ့ item ကို နှိပ်ပါ။ Auto-completion control ထဲက item တစ်ခုကို နှိပ်လိုက်တာက ၎င်းကို လက်ရှိ cursor ရှိနေတဲ့ နေရာမှာ ထည့်သွင်း (insert) ပေးသွားမှာ ဖြစ်ပါတယ်။

ပုံမှန်အားဖြင့် (By default) auto-completion control က သင် စာရိုက်နေစဉ်မှာပဲ အသုံးပြုနိုင်တဲ့ terms စာရင်းကို အကြံပြု ပြသပေးမှာပါ။ ဥပမာ၊ သင်က bar chart နဲ့အတူ pie chart ကိုပါ ထည့်ချင်တယ် ဆိုပါစို့။ Pie chart ကို သီးခြား subcommand တစ်ခုနဲ့ သတ်မှတ်ပေးရပါတယ်။

၂။ FREQ keyword ရဲ့ နောက်မှာ Enter နှိပ်ပြီး၊ subcommand တစ်ခုရဲ့ အစကို ညွှန်ပြဖို့အတွက် forward slash (/) တစ်ခုကို ရိုက်ထည့်ပါ။ Syntax Editor က လက်ရှိ command အတွက် subcommands စာရင်းကို အကြံပြု ပြသပေးပါလိမ့်မယ်။

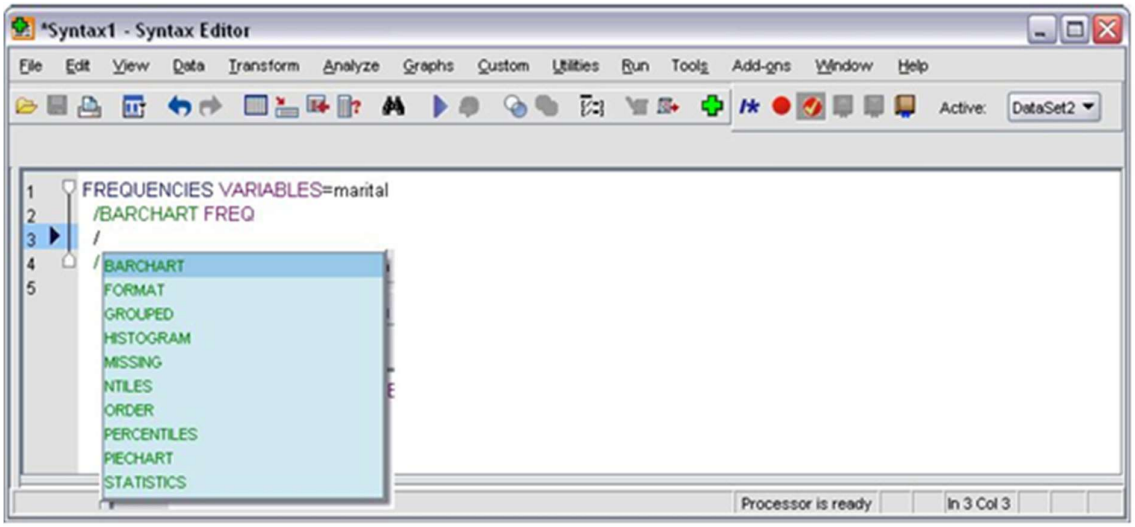


Figure 71. Auto-completion control displaying subcommands

လက်ရှိ command အတွက် ပိုမို အသေးစိတ်ကျတဲ့ help ကို ရယူချင်ရင် F1 key ကို နှိပ်ပါ။ ဒါက သင့်ကို လက်ရှိ command အတွက် command syntax reference information ဆီ တိုက်ရိုက် ခေါ်ဆောင်သွားမှာ ဖြစ်ပါတယ်။

Syntax window ထဲမှာ ပြသထားတဲ့ text တွေမှာ အရောင်လေးတွေ ပါနေတာကို သင် သတိထားမိပါလိမ့်မယ်။ အသိအမှတ်ပြုထားတဲ့ (recognized) terms တွေကိုပဲ

အရောင်ချယ်ပေးထားတဲ့အတွက်၊ color coding က အသိအမှတ်မပြုထားတဲ့ (unrecognized) terms တွေကို မြန်မြန်ဆန်ဆန် ခွဲခြားသိမြင်ဖို့ ကူညီပေးပါတယ်။ ဥပမာအနေနဲ့ သင်က FORMAT subcommand ကို FRMAT ဆိုပြီး စာလုံးပေါင်း မှားရိုက်မိတယ် ဆိုပါစို့။ Subcommands တွေကို ပုံမှန်အားဖြင့် အစိမ်းရောင် ချယ်ပေးထားပေမယ့်၊ FRMAT ဆိုတဲ့ text ကိုတော့ unrecognized ဖြစ်တဲ့အတွက် အရောင်မပါဘဲ ပြသပေးနေမှာ ဖြစ်ပါတယ်။

Opening and Running a Syntax File

၁။ Save လုပ်ထားတဲ့ syntax file တစ်ခုကို ဖွင့်ဖို့၊ menus တွေထဲကနေ File > Open > Syntax... ကို ရွေးချယ်ပါ။ ဖိုင်တွေဖွင့်ဖို့အတွက် standard dialog box တစ်ခု ပေါ်လာပါလိမ့်မယ်။

၂။ syntax file တစ်ခုကို ရွေးချယ်ပါ။ တကယ်လို့ syntax files တွေကို မမြင်ရဘူးဆိုရင်၊ သင်ကြည့်ချင်တဲ့ file type အနေနဲ့ Syntax (*.sps) ကို ရွေးချယ်ထားခြင်း ရှိမရှိ သေချာစစ်ဆေးပါ။

၃။ Open ကို နှိပ်ပါ။

၄။ Commands တွေကို run ဖို့ syntax window ထဲက Run menu ကို အသုံးပြုပါ။

တကယ်လို့ commands တွေက သတ်မှတ်ထားတဲ့ data file တစ်ခုအတွက် သုံးတာဖြစ်နေမယ်ဆိုရင်၊ commands တွေကို မ run ခင် အဲဒီ data file ကို အရင်ဖွင့်ထားရပါမယ်၊ ဒါမှမဟုတ် data file ကို ဖွင့်ပေးမယ့် command တစ်ခု ပါဝင်ရပါမယ်။ Data files တွေကို ဖွင့်ပေးတဲ့ dialog boxes တွေကနေ ဒီလို command မျိုးကို သင် paste လုပ်ယူနိုင်ပါတယ်။

Using Breakpoints

Breakpoints တွေက သင့်ကို syntax အတွင်း သင်သတ်မှတ်ထားတဲ့ နေရာတွေမှာ command syntax ရဲ့ execution (လုပ်ဆောင်မှု) ကို ရပ်တန့်စေပြီး၊ အဆင်သင့်ဖြစ်တဲ့အခါမှ execution ကို ဆက်လက်လုပ်ဆောင်ခွင့် ပေးပါတယ်။ ဒါက သင့်ကို syntax job ရဲ့ ကြားကာလ အဆင့်တစ်ခုခုမှာ output ဒါမှမဟုတ် data တွေကို ကြည့်ရှုခွင့် ပေးသလို၊ FREQUENCIES လိုမျိုး data ရဲ့ လက်ရှိ အခြေအနေနဲ့ ပတ်သက်တဲ့ သတင်းအချက်အလက်ကို ပြသပေးမယ့် command syntax မျိုးကို run ဖို့လည်း ခွင့်ပြုပါတယ်။ Breakpoints တွေကို command တစ်ခုရဲ့ level မှာပဲ သတ်မှတ်လို့ရပြီး၊ command တစ်ခုအတွင်းက သီးခြား lines တွေအပေါ်မှာတော့ သတ်မှတ်လို့ မရပါဘူး။

Command တစ်ခုပေါ်မှာ breakpoint တစ်ခု ထည့်သွင်း (insert) ဖို့အတွက်-

၁။ Command နဲ့ သက်ဆိုင်တဲ့ text ရဲ့ ဘယ်ဘက် ဇရိယာထဲက ဘယ်နေရာမှာမဆို ကလစ်နှိပ်ပါ။

သင် ဘယ်နေရာကိုပဲ နှိပ်ခဲ့သည်ဖြစ်စေ breakpoint ကို command text ရဲ့ ဘယ်ဘက် ဇရိယာထဲက command name နဲ့ တစ်ကြောင်းတည်းပေါ်မှာ အနီရောင် စက်ဝိုင်းလေးတစ်ခု အနေနဲ့ ကိုယ်စားပြု ပြသပေးမှာ ဖြစ်ပါတယ်။

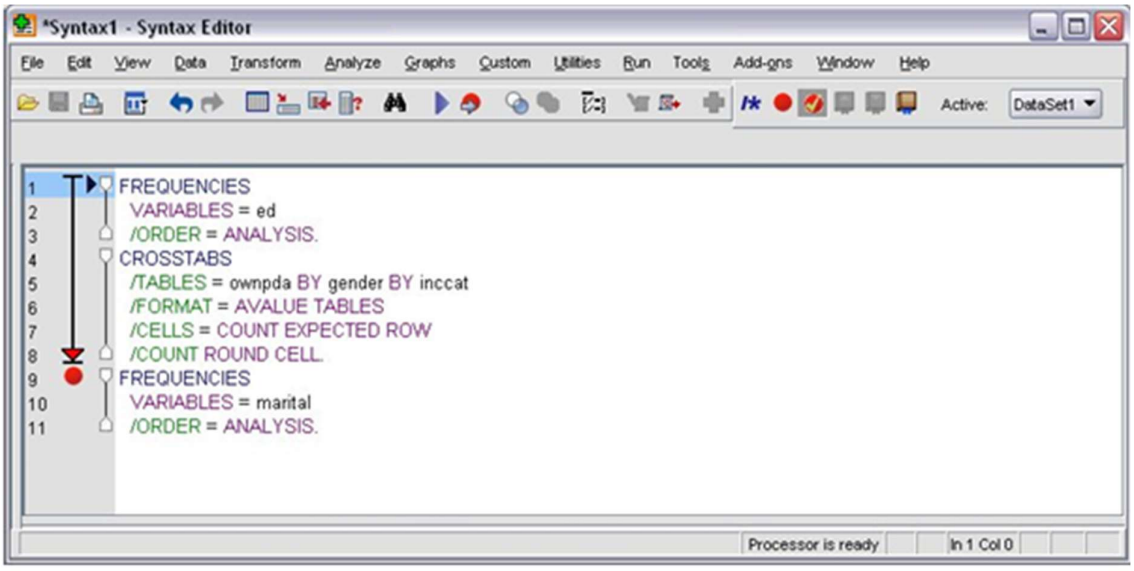

*Syntax1 - Syntax Editor
File Edit View Data Transform Analyze Graphs Custom Utilities Run Tools Add-ons Window Help
Active: DataSet1
1 FREQUENCIES
2 VARIABLES = ed
3 /ORDER = ANALYSIS.
4 CROSSTABS
5 /TABLES = ownpda BY gender BY inccat
6 /FORMAT = AVALUE TABLES
7 /CELLS = COUNT EXPECTED ROW
8 /COUNT ROUND CELL.
9 FREQUENCIES
10 VARIABLES = marital
11 /ORDER = ANALYSIS.
Processor is ready
In 1 Col 0

Figure 72. Execution stopped at a breakpoint

Breakpoints တွေ ပါဝင်နေတဲ့ command syntax ကို သင် run တဲ့အခါ၊ breakpoint ပါဝင်နေတဲ့ command တစ်ခုစီရဲ့ အရှေ့မှာ execution က ရပ်တန့်သွားမှာ ဖြစ်ပါတယ်။

Command text ရဲ့ ဘယ်ဘက်က အောက်စိုက်နေတဲ့ မြှား (downward pointing arrow) လေးက syntax run ရဲ့ တိုးတက်မှု (progress) ကို ပြသပေးပါတယ်။ သူက ပထမဆုံး run တဲ့ command ကနေစပြီး နောက်ဆုံး run တဲ့ command အထိ ဇရိယာတစ်ခုလုံးကို ဖြန့်ကျက်ပြသပေးတဲ့အတွက်၊ breakpoints တွေ ပါဝင်တဲ့ command syntax ကို run တဲ့အခါ အထူးသဖြင့် အသုံးဝင်ပါတယ်။

Breakpoint ရဲ့နောက်ကနေ execution ကို ဆက်လက်လုပ်ဆောင် (resume) ဖို့အတွက်-

၂။ Syntax Editor ရဲ့ menus တွေထဲကနေ Run > Continue ကို ရွေးချယ်ပါ။

Chapter 8. Modifying Data Values

သင် စတင်အသုံးပြုမယ့် data တွေက သင့်ရဲ့ analysis ဒါမှမဟုတ် reporting လိုအပ်ချက်တွေအတွက် အမြဲတမ်း အသုံးဝင်ဆုံးဖြစ်အောင် ဖွဲ့စည်းထားတာမျိုး မဟုတ်နိုင်ပါဘူး။ ဥပမာအားဖြင့်၊ သင်က အောက်ပါအရာတွေကို လုပ်ချင်နိုင်ပါတယ်-

- scale variable တစ်ခုကနေ categorical variable တစ်ခု ဖန်တီးချင်တာမျိုး။
- response categories အများကြီးကို category တစ်ခုတည်းအဖြစ် ပေါင်းစည်းချင်တာမျိုး။
- ရှိပြီးသား variables နှစ်ခုရဲ့ တွက်ချက်ထားတဲ့ ကွာခြားချက် (computed difference) ဖြစ်တဲ့ variable အသစ်တစ်ခုကို ဖန်တီးချင်တာမျိုး။
- dates နှစ်ခုကြားက အချိန်ကာလ (length of time) ကို တွက်ချက်ချင်တာမျိုးတွေ ဖြစ်နိုင်ပါတယ်။

ဒီ chapter မှာတော့ demo.sav ဆိုတဲ့ data file ကို သုံးသွားမှာပါ။ ပိုပြီးသိချင်ရင်တော့ Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ သွားကြည့်နိုင်ပါတယ်။

Creating a Categorical Variable from a Scale Variable

demo.sav data file ထဲမှာရှိတဲ့ categorical variables တော်တော်များများက တကယ်တော့ အဲဒီ data file ထဲက scale variables တွေကနေ ဆင်းသက်လာတာပါ။ ဥပမာအနေနဲ့ inccat ဆိုတဲ့ variable က income ကို categories လေးခုအဖြစ် ခွဲခြားထားတာသက်သက်ပါပဲ။ ဒီ categorical variable ဟာ အောက်ပါ income categories (ထောင်ဂဏန်းနဲ့) တွေကို ကိုယ်စားပြုဖို့အတွက် 1 ကနေ 4 အထိ integer values တွေကို အသုံးပြုထားပါတယ်- $25 အောက်၊ $25–$49၊ $50–$74 နဲ့ $75 သို့မဟုတ် အဲဒီထက်ပိုတဲ့ ပမာဏတွေပဲ ဖြစ်ပါတယ်။

inccat ဆိုတဲ့ categorical variable ကို ဖန်တီးဖို့အတွက်-

၁။ Data Editor window ရဲ့ menus တွေထဲကနေ Transform > Visual Binning... ကို ရွေးချယ်ပါ။

ပထမဆုံးပေါ်လာမယ့် Visual Binning dialog box ထဲမှာ၊ binned variables အသစ်တွေ ဖန်တီးချင်တဲ့ scale ဒါမှမဟုတ် ordinal variables တွေကို ရွေးချယ်ပေးရပါမယ်။ Binning ဆိုတာက ဆက်တိုက်ဖြစ်နေတဲ့ (contiguous) values နှစ်ခု ဒါမှမဟုတ် အဲဒီထက်ပိုတာတွေကို ယူပြီး category တစ်ခုတည်းအောက်မှာ စုစည်း (grouping) လိုက်တာကို ဆိုလိုတာပါ။

Visual Binning က ကောင်းမွန်တဲ့ binning choices တွေ လုပ်နိုင်ဖို့အတွက် data file ထဲက အမှန်တကယ် values တွေအပေါ် အားကိုးရတဲ့အတွက်၊ သူက data file ကို အရင်ဆုံး ဖတ်ဖို့ လိုအပ်ပါတယ်။ တကယ်လို့ သင့်ရဲ့ data file ထဲမှာ cases အရေအတွက် အများကြီး ပါဝင်နေမယ်ဆိုရင် အချိန်ယူရနိုင်တဲ့အတွက်၊ ဒီ initial dialog box မှာ ဖတ်မယ့် (scan လုပ်မယ့်) cases အရေအတွက်ကို ကန့်သတ်ထားလို့ရအောင်လည်း ခွင့်ပြုပေးထားပါတယ်။ ကျွန်တော်တို့ရဲ့ sample data file အတွက်တော့ ဒါက မလိုအပ်ပါဘူး။ သူ့ထဲမှာ cases ပေါင်း ၆၀၀၀ ကျော် ပါဝင်နေပေမယ့်လည်း၊ အဲဒီ cases အရေအတွက်ကို scan လုပ်ဖို့ အချိန်အများကြီး မကြာပါဘူး။

၂။ Variables list ထဲကနေ Household income in thousands [income] ကို Variables to Bin list ထဲ ဆွဲထည့် (drag and drop) ပါ။ ပြီးရင် Continue ကို နှိပ်ပါ။

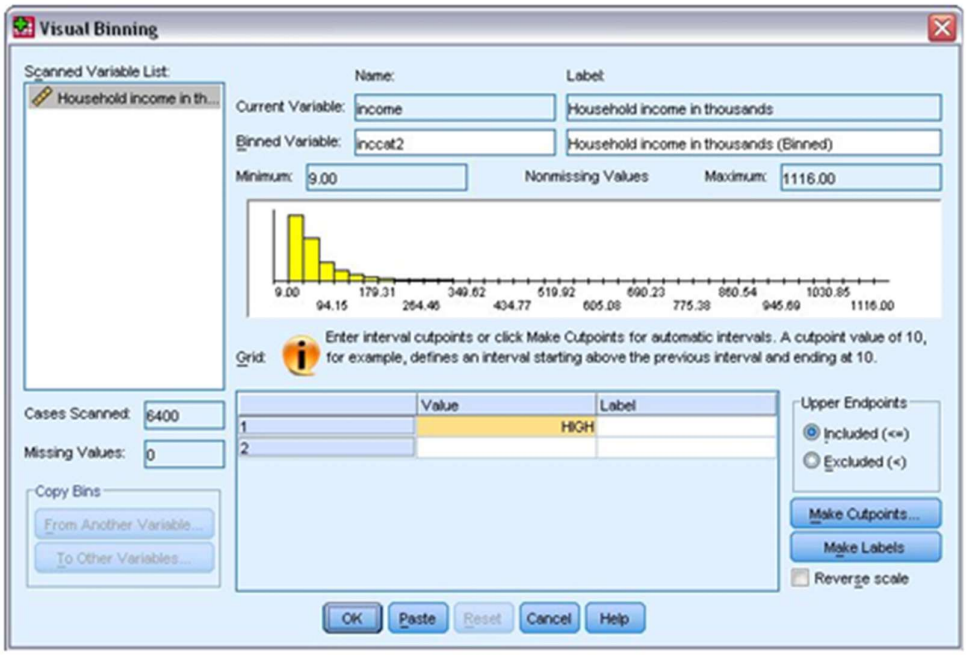


Figure 73. Main Visual Binning dialog box

၃။ Main Visual Binning dialog box ထဲက Scanned Variable List မှာ Household income in thousands [income] ကို ရွေးချယ်ပါ။ (ဒီကိစ္စမှာ အရမ်းကို skewed ဖြစ်နေတဲ့) ရွေးချယ်ထားတဲ့ variable ရဲ့ distribution ကို histogram တစ်ခုက ပြသပေးပါလိမ့်မယ်။

၄။ binned variable အသစ်ရဲ့ နာမည်အတွက် inccat2 လို့ ရိုက်ထည့်ပြီး၊ variable label အတွက် Income category [in thousands] လို့ ရိုက်ထည့်ပါ။

၅။ Make Cutpoints ကို နှိပ်ပါ။

၆။ Equal Width Intervals ကို ရွေးချယ်ပါ။

၇။ ပထမဆုံး cutpoint location အတွက် 25၊ number of cutpoints အတွက် 3 နဲ့ width အတွက် 25 လို့ ရိုက်ထည့်ပါ။

Binned categories အရေအတွက်ဟာ cutpoints အရေအတွက်ထက် တစ်ခု ပိုများပါတယ်။ ဒါကြောင့် ဒီဥပမာမှာဆိုရင် binned variable အသစ်မှာ categories လေးခု ရှိလာမှာဖြစ်ပြီး၊ ပထမဆုံး categories သုံးခုစီမှာ 25 (ထောင်ဂဏန်း) ranges တွေ ပါဝင်မှာဖြစ်ကာ၊ နောက်ဆုံးတစ်ခုမှာတော့ အမြင့်ဆုံး cutpoint value ဖြစ်တဲ့ 75 (ထောင်ဂဏန်း) အထက်က values တွေအားလုံး ပါဝင်မှာ ဖြစ်ပါတယ်။

၈။ Apply ကို နှိပ်ပါ။

အခု grid ထဲမှာ ပြသနေတဲ့ values တွေဟာ သတ်မှတ်ထားတဲ့ cutpoints တွေကို ကိုယ်စားပြုပြီး၊ အဲဒါတွေက category တစ်ခုစီရဲ့ upper endpoints တွေပဲ ဖြစ်ပါတယ်။ Histogram ထဲက ဒေါင်လိုက်မျဉ်း (vertical lines) တွေကလည်း cutpoints တွေရဲ့ နေရာတွေကို ညွှန်ပြနေပါတယ်။

ပုံမှန်အားဖြင့် (By default) ဒီ cutpoint values တွေကို သက်ဆိုင်ရာ categories တွေထဲမှာ ထည့်သွင်း (included) ပေးထားပါတယ်။ ဥပမာအနေနဲ့၊ ပထမဆုံး value ဖြစ်တဲ့ 25 ဟာ 25 ထက်နည်းတဲ့ ဒါမှမဟုတ် 25 နဲ့ညီမျှတဲ့ values အားလုံးကို ပါဝင်စေမှာပါ။ ဒါပေမယ့် ဒီဥပမာမှာတော့ ကျွန်တော်တို့က 25 အောက်၊ 25–49၊ 50–74 နဲ့ 75 သို့မဟုတ် အဲဒီထက်ပိုတဲ့ ပမာဏတွေနဲ့ ကိုက်ညီတဲ့ categories တွေကို လိုချင်တာပါ။

၉။ Upper Endpoints group ထဲမှာ Excluded (<) ကို ရွေးချယ်ပါ။

၁၀။ ပြီးရင် Make Labels ကို နှိပ်ပါ။

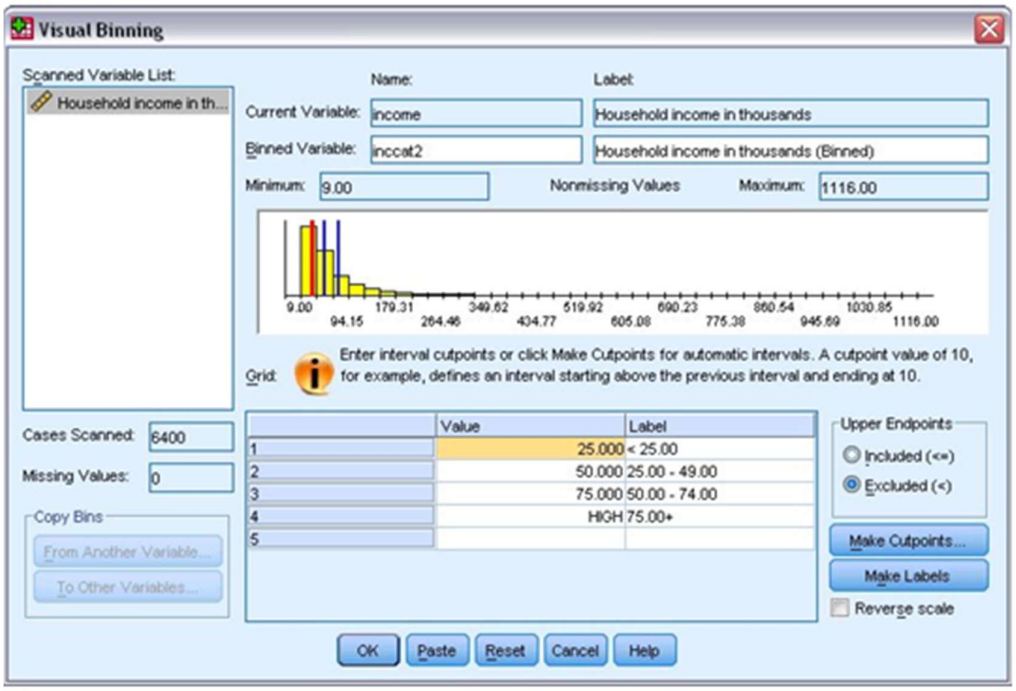


Figure 74. Automatically generated value labels

ဒါက category တစ်ခုစီအတွက် descriptive value labels တွေကို အလိုအလျောက် ဖန်တီးပေးသွားမှာပါ။ Binned variable အသစ်အတွက် သတ်မှတ်ပေးထားတဲ့ အမှန်တကယ် values တွေက 1 ကနေစတဲ့ အစဉ်လိုက် integers တွေသာ ဖြစ်တဲ့အတွက်၊ ဒီ value labels တွေက အလွန် အသုံးဝင်နိုင်ပါတယ်။

သင်က grid ထဲမှာ cutpoints တွေနဲ့ labels တွေကို ကိုယ်တိုင် (manually) ရိုက်ထည့်တာ ဒါမှမဟုတ် ပြောင်းလဲတာတွေ လုပ်နိုင်သလို၊ histogram ထဲက cutpoint လိုင်းတွေကို ဆွဲရွှေ့ (dragging and dropping) ခြင်းအားဖြင့် cutpoint နေရာတွေကို ပြောင်းလဲတာ၊ cutpoint လိုင်းတွေကို histogram အပြင်ဘက်ကို ဆွဲထုတ်ခြင်းအားဖြင့် cutpoints တွေကို ဖျက်ပစ် (delete) တာတွေကိုလည်း လုပ်ဆောင်နိုင်ပါသေးတယ်။

၁၁။ binned variable အသစ်ကို ဖန်တီးဖို့ OK ကို နှိပ်ပါ။ Variable အသစ်ကို Data Editor ထဲမှာ ပြသပေးပါလိမ့်မယ်။ Variable ကို ဖိုင်ရဲ့ နောက်ဆုံးမှာ ထည့်သွင်းလိုက်တာ ဖြစ်တဲ့အတွက်၊ သူ့ကို Data View မှာဆိုရင် ညာဘက်အစွန်ဆုံး column မှာ ပြသပြီး၊ Variable View မှာဆိုရင်တော့ နောက်ဆုံး row မှာ ပြသပေးမှာ ဖြစ်ပါတယ်။

Computing New Variables

သခ်ၤာဆိုင်ရာ functions မ်ိုးစုံကို အသုံးပြုပြီး၊ အလွန် ရှုပ်ထွေးတဲ့ ညီမျှခြင်း (equations) တွေအပေါ် အခြေခံကာ variables အသစ်တွေကို တွက်ချက် (compute) နိုင်ပါတယ်။ ဒါပေမယ့် ဒီဥပမာမှာတော့၊ ရှိပြီးသား variables နှစ်ခုရဲ့ values တွေကြားက ကွာခြားချက်ဖြစ်တဲ့ variable အသစ်တစ်ခုကို ရိုးရှင်းစွာပဲ တွက်ချက်ကြည့်ကြပါမယ်။

demo.sav data file ထဲမှာ ဖြေဆိုသူရဲ့ လက်ရှိအသက် (current age) အတွက် variable တစ်ခုနဲ့ လက်ရှိအလုပ်မှာ လုပ်သက်ဘယ်နှစ်နှစ်ရှိပြီလဲ (number of years at current job) ဆိုတဲ့ variable တစ်ခု ပါဝင်ပါတယ်။ ဒါပေမယ့် အဲဒီအလုပ်ကို စတင်လုပ်ကိုင်ခဲ့ချိန်က ဖြေဆိုသူရဲ့အသက်ကိုပြတဲ့ variable တော့ မပါဝင်ပါဘူး။ လက်ရှိအသက်နဲ့ လက်ရှိအလုပ်မှာလုပ်ခဲ့တဲ့ နှစ်အရေအတွက်တို့ကြားက တွက်ချက်ထားတဲ့ ကွာခြားချက်ဖြစ်တဲ့ variable အသစ်တစ်ခုကို ကျွန်တော်တို့ ဖန်တီးနိုင်ပါတယ်။ အဲဒါက ဖြေဆိုသူ အဲဒီအလုပ်ကို စတင်လုပ်ကိုင်ခဲ့တဲ့ ခန့်မှန်းခြေအသက် (approximate age) ပဲ ဖြစ်ပါလိမ့်မယ်။

၁။ Data Editor window ရဲ့ menus တွေထဲကနေ Transform > Compute Variable... ကို ရွေးချယ်ပါ။

၂။ Target Variable အတွက် jobstart လို့ ရိုက်ထည့်ပါ။

၃။ source variable list ထဲမှာ Age in years [age] ကို ရွေးချယ်ပြီး၊ သူ့ကို Numeric Expression text box ထဲ copy ကူးထည့်ဖို့ မြှားခလုတ် (arrow button) လေးကို နှိပ်ပါ။

၄။ dialog box ထဲက calculator pad ပေါ်က အနှုတ်လက္ခဏာ (–) ခလုတ်ကို နှိပ်ပါ (ဒါမှမဟုတ် keyboard ပေါ်က အနှုတ်ခလုတ်ကို နှိပ်ပါ)။

၅။ Years with current employer [employ] ကို ရွေးချယ်ပြီး၊ သူ့ကို expression ထဲ copy ကူးထည့်ဖို့ မြှားခလုတ်လေးကို နှိပ်ပါ။

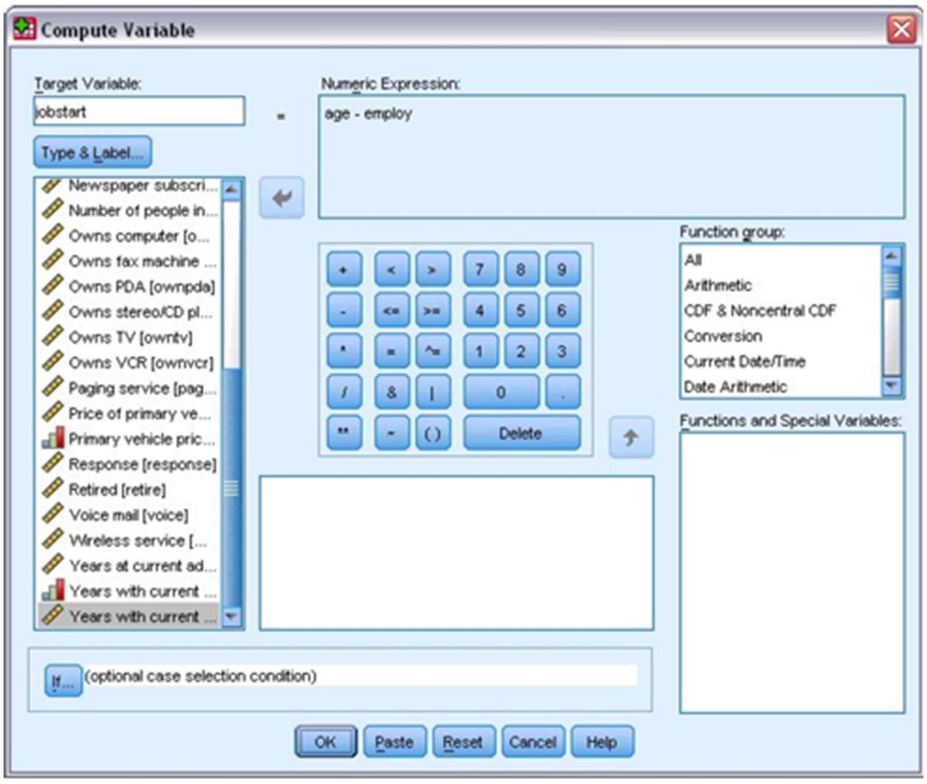


Figure 75. Compute Variable dialog box

မှတ်ချက် - မှန်ကန်တဲ့ employment variable ကို ရွေးချယ်မိဖို့ သတိပြုပါ။ သင်မလိုချင်တဲ့ recoded categorical version ဖြစ်တဲ့ variable လည်း ရှိနေတတ်ပါတယ်။ Numeric expression ဟာ age–employ သာ ဖြစ်သင့်ပြီး၊ age–empcat မဖြစ်သင့်ပါဘူး။

၆။ Variable အသစ်ကို တွက်ချက်ဖို့ OK ကို နှိပ်ပါ။

Variable အသစ်ကို Data Editor ထဲမှာ ပြသပေးပါလိမ့်မယ်။ Variable ကို ဖိုင်ရဲ့ နောက်ဆုံးမှာ ထည့်သွင်းလိုက်တာ ဖြစ်တဲ့အတွက်၊ သူ့ကို Data View မှာဆိုရင် ညာဘက်အစွန်ဆုံး column မှာ ပြသပြီး၊ Variable View မှာဆိုရင်တော့ နောက်ဆုံး row မှာ ပြသပေးမှာ ဖြစ်ပါတယ်။

Using Functions in Expressions

Expressions တွေထဲမှာ ကြိုတင်သတ်မှတ်ထားတဲ့ (predefined) functions တွေကိုလည်း အသုံးပြုနိုင်ပါတယ်။ အောက်ပါတို့အပါအဝင် built-in functions ၇၀ ကျော်ကို အသုံးပြုနိုင်ပါတယ်-

- Arithmetic functions
- Statistical functions

- Distribution functions
- Logical functions
- Date and time aggregation and extraction functions
- Missing-value functions
- Cross-case functions
- String functions

Functions တွေကို arithmetic operations အတွက် group တစ်ခု၊ statistical metrics တွေ တွက်ချက်ဖို့အတွက် နောက်ထပ် group တစ်ခု စသဖြင့် ယုတ္တိတန်တဲ့ သီးခြား groups တွေအဖြစ် ဖွဲ့စည်းထားပါတယ်။ ပိုပြီး အဆင်ပြေစေဖို့အတွက် $TIME (လက်ရှိရက်စွဲနဲ့ အချိန်) လိုမျိုး အသုံးများတဲ့ system variables အတော်များများကိုလည်း သင့်လျော်တဲ့ function groups တွေထဲမှာ ထည့်သွင်းပေးထားပါတယ်။

Pasting a Function into an Expression

Function တစ်ခုကို expression ထဲကို paste လုပ်ထည့်ဖို့အတွက်-

၁။ သင် function ကို ပေါ်စေချင်တဲ့ expression ထဲက နေရာမှာ cursor ကို ချထားပါ။

၂။ Function group list ထဲကနေ သင့်လျော်တဲ့ group ကို ရွေးချယ်ပါ။ All လို့ နာမည်တပ်ထားတဲ့ group ကတော့ အသုံးပြုနိုင်တဲ့ functions နဲ့ system variables အားလုံးရဲ့ စာရင်းကို ပြသပေးမှာပါ။

၃။ Functions and Special Variables list ထဲက function ကို double-click နှိပ်ပါ (ဒါမှမဟုတ် function ကို ရွေးချယ်ပြီး Function group list ရဲ့ဘေးက မြှားလေးကို နှိပ်ပါ)။

Function ကို expression ထဲမှာ ထည့်သွင်း (inserted) ပေးသွားပါလိမ့်မယ်။ တကယ်လို့ သင်က expression ရဲ့ တစ်စိတ်တစ်ပိုင်းကို highlight လုပ်ထားပြီးမှ function ကို ထည့်သွင်းမယ်ဆိုရင်၊ expression ရဲ့ highlight လုပ်ထားတဲ့ အပိုင်းကို function ထဲက ပထမဆုံး argument အနေနဲ့ အသုံးပြုသွားမှာ ဖြစ်ပါတယ်။

Editing a Function in an Expression

Paste လုပ်လိုက်တဲ့ function ထဲမှာ မေးခွန်းသင်္ကေတ (question marks) တွေနဲ့ ကိုယ်စားပြုထားတဲ့ arguments တွေကို သင် မရိုက်ထည့်မချင်း function ဟာ ပြည့်စုံမှာ မဟုတ်ပါဘူး။ မေးခွန်းသင်္ကေတ အရေအတွက်ဟာ function တစ်ခု ပြည့်စုံဖို့အတွက် လိုအပ်တဲ့ အနည်းဆုံး arguments အရေအတွက်ကို ညွှန်ပြနေတာပါ။

၁။ Paste လုပ်ထားတဲ့ function ထဲက မေးခွန်းသင်္ကေတ (question marks) တွေကို highlight လုပ်ပါ။

၂။ Arguments တွေကို ရိုက်ထည့်ပါ။ တကယ်လို့ arguments တွေက variable names တွေ ဖြစ်နေမယ်ဆိုရင်၊ variable list ထဲကနေ paste လုပ်ယူနိုင်ပါတယ်။

Using Conditional Expressions

ရွေးချယ်ထားတဲ့ cases တွေရဲ့ subsets တွေအပေါ် transformations တွေ လုပ်ဆောင်ဖို့အတွက် conditional expressions (logical expressions လို့လည်း ခေါ်ပါတယ်) တွေကို အသုံးပြုနိုင်ပါတယ်။ Conditional expression တစ်ခုဟာ case တစ်ခုချင်းစီအတွက် true, false ဒါမှမဟုတ် missing ဆိုတဲ့ value တစ်ခုကို ပြန်ပေး (returns) ပါတယ်။ တကယ်လို့ conditional expression ရဲ့ ရလဒ်က true ဖြစ်တယ်ဆိုရင်၊ အဲဒီ case အပေါ်မှာ transformation ကို လုပ်ဆောင်ပါလိမ့်မယ်။ ရလဒ်က false ဒါမှမဟုတ် missing သာ ဖြစ်နေမယ်ဆိုရင်တော့၊ အဲဒီ case ပေါ်မှာ transformation ကို လုပ်ဆောင်မှာ မဟုတ်ပါဘူး။

Conditional expression တစ်ခုကို သတ်မှတ်ဖို့အတွက်-

၁။ Compute Variable dialog box ထဲက If ကို နှိပ်ပါ။ ဒါက If Cases dialog box ကို ဖွင့်ပေးပါလိမ့်မယ်။

၂။ Include if case satisfies condition ကို ရွေးချယ်ပါ။

၃။ Conditional expression ကို ရိုက်ထည့်ပါ။

Conditional expressions အများစုမှာ အောက်ပါအတိုင်း အနည်းဆုံး relational operator တစ်ခု ပါဝင်တတ်ပါတယ်-

age>=21

သို့မဟုတ်

income*3<100

ပထမ ဥပမာမှာဆိုရင် Age [age] အတွက် value 21 ဒါမှမဟုတ် အဲဒီထက်ပိုတဲ့ cases တွေကိုသာ ရွေးချယ်မှာပါ။ ဒုတိယ ဥပမာမှာတော့ case တစ်ခုကို ရွေးချယ်ခံရဖို့အတွက် Household income in thousands [income] ကို 3 နဲ့ မြှောက်ထားတဲ့ရလဒ်ဟာ 100 ထက် နည်းနေဖို့ လိုအပ်ပါတယ်။

Logical operators တွေကို အသုံးပြုပြီးတော့လည်း conditional expressions နှစ်ခု ဒါမှမဟုတ် အဲဒီထက်ပိုတာတွေကို အောက်ပါအတိုင်း ချိတ်ဆက် (link) နိုင်ပါသေးတယ်- age>=21 | ed>=4 သို့မဟုတ် income*3<100 & ed=5

ပထမ ဥပမာမှာဆိုရင် Age [age] condition ဒါမှမဟုတ် Level of education [ed] condition တစ်ခုခုနဲ့ ကိုက်ညီတဲ့ cases တွေကို ရွေးချယ်မှာပါ။ ဒုတိယ ဥပမာမှာဆိုရင်တော့ case တစ်ခု ရွေးချယ်ခံရဖို့အတွက် Household income in thousands [income] ရော Level of education [ed] conditions နှစ်ခုလုံးနဲ့ ကိုက်ညီနေဖို့ လိုအပ်ပါတယ်။

Working with Dates and Times

Dates တွေ times တွေနဲ့ ပတ်သက်ပြီး အများဆုံး လုပ်လေ့ရှိတဲ့ tasks အတော်များများကို Date and Time Wizard သုံးပြီး အလွယ်တကူ လုပ်ဆောင်နိုင်ပါတယ်။ ဒီ wizard ကို အသုံးပြုပြီး သင်ဟာ-

- Date ဒါမှမဟုတ် time ပါဝင်တဲ့ string variable တစ်ခုကနေ date/time variable တစ်ခု ဖန်တီးတာ။
- Date ဒါမှမဟုတ် time ရဲ့ မတူညီတဲ့ အစိတ်အပိုင်းတွေ ပါဝင်နေတဲ့ variables တွေကို ပေါင်းစပ်ပြီး (merging) date/time variable တစ်ခု တည်ဆောက်တာ။
- Date/time variables တွေကနေ values တွေကို ပေါင်းတာ ဒါမှမဟုတ် နှုတ်တာ (date/time variables နှစ်ခုကို အချင်းချင်း ပေါင်းတာ နှုတ်တာ အပါအဝင်) ။
- Date ဒါမှမဟုတ် time variable တစ်ခုဆီက အစိတ်အပိုင်းတစ်ခုကို ထုတ်ယူတာ (ဥပမာ၊ mm/dd/yyyy ပုံစံရှိတဲ့ date/time variable ကနေ လတစ်လရဲ့ ရက် (day of month) ကို ထုတ်ယူတာ) တွေကို လုပ်ဆောင်နိုင်ပါတယ်။

ဒီ အပိုင်းမှာပါတဲ့ ဥပမာတွေကတော့ upgrade.sav ဆိုတဲ့ data file ကို အသုံးပြုထားတာပါ။ ပိုပြီး သိချင်တယ်ဆိုရင် Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ သွားကြည့်နိုင်ပါတယ်။

Date and Time Wizard ကို အသုံးပြုဖို့-

၁။ menus တွေထဲကနေ Transform > Date and Time Wizard... ကို ရွေးချယ်ပါ။

Date and Time Wizard
Welcome to the date and time wizard
What would you like to do?
Learn how dates and times are represented
Create a date/time variable from a string containing a date or time
Create a date/time variable from variables holding parts of dates or times
Calculate with dates and times
Extract a part of a date or time variable
Assign periodicity to a dataset (for time series data). This ends the wizard and opens the Define Dates dialog box
< Back
Next >
Finish
Cancel
Help

Figure 76. Date and Time Wizard introduction screen

Date and Time Wizard ရဲ့ ပထမဆုံး မိတ်ဆက် မျက်နှာပြင် (introduction screen) က သင့်ကို အထွေထွေ tasks လေးတွေ ပြသပေးမှာပါ။ လက်ရှိ data နဲ့ မသက်ဆိုင်တဲ့ tasks တွေကိုတော့ ပိတ်ထား (disabled) ပါလိမ့်မယ်။ ဥပမာအနေနဲ့၊ upgrade.sav ဆိုတဲ့ data file ထဲမှာ ဘယ် string variables မှ မပါဝင်တဲ့အတွက်၊ string ကနေ date variable တစ်ခု ဖန်တီးဖို့ဆိုတဲ့ task ကို ပိတ်ထားမှာ ဖြစ်ပါတယ်။

တကယ်လို့ သင်က IBM SPSS Statistics ထဲက dates နဲ့ times တွေအကြောင်း အသစ်ဖြစ်နေသေးတယ် ဆိုရင်တော့၊ Learn how dates and times are represented ကို ရွေးချယ်ပြီး Next ကို နှိပ်နိုင်ပါတယ်။ ဒါက date/time variables တွေအကြောင်း အကျဉ်းချုပ် ရှင်းလင်းချက်လေးတွေ ပြသပေးမယ့် မျက်နှာပြင်ဆီ ခေါ်ဆောင်သွားမှာဖြစ်ပြီး၊ Help ခလုတ်ကတစ်ဆင့် ပိုမို အသေးစိတ်ကျတဲ့ သတင်းအချက်အလက်တွေဆီ သွားနိုင်မယ့် link တစ်ခုလည်း ပါဝင်ပါတယ်။

Calculating the Length of Time between Two Dates

Dates တွေနဲ့ ပတ်သက်ရင် အလုပ်အများဆုံး tasks တွေထဲက တစ်ခုကတော့ dates နှစ်ခုကြားက အချိန်ကာလ (length of time) ကို တွက်ချက်တာပါပဲ။ ဥပမာအနေနဲ့၊ customer တစ်ယောက်ချင်းစီက နောက်ဆုံး upgrade ဝယ်ယူခဲ့တဲ့ အချိန်ကနေစပြီး ဘယ်နှစ်နှစ် ကြာသွားပြီလဲဆိုတာကို ရှာဖွေခြင်းအားဖြင့်၊ upgrade လိုင်စင် ဝယ်ယူမှုတွေကို analyze လုပ်ဖို့ စိတ်ဝင်စားနေတဲ့ software ကုမ္ပဏီတစ်ခုကို စဉ်းစားကြည့်ပါ။ upgrade.sav data file ထဲမှာ customer တစ်ယောက်ချင်းစီ နောက်ဆုံး upgrade ဝယ်ယူခဲ့တဲ့ ရက်စွဲအတွက် variable တစ်ခု ပါဝင်ပေမယ့်၊ အဲဒီ ဝယ်ယူမှုကနေစပြီး ဘယ်နှစ်နှစ်ကြာပြီလဲ ဆိုတာတော့ မပါပါဘူး။ နောက်ဆုံး upgrade လုပ်ခဲ့တဲ့ ရက်စွဲနဲ့ နောက်ထပ် product ထွက်မယ့် (release) ရက်စွဲတို့ကြားက အချိန်ကာလကို နှစ်နဲ့တွက်ထားတဲ့ variable အသစ်တစ်ခုက ဒီပမာဏကို တိုင်းတာပေးနိုင်ပါလိမ့်မယ်။

Dates နှစ်ခုကြားက အချိန်ကာလကို တွက်ချက်ဖို့အတွက်-

၁။ Date and Time Wizard ရဲ့ မိတ်ဆက် မျက်နှာပြင်မှာ Calculate with dates and times ကို ရွေးချယ်ပြီး Next ကို နှိပ်ပါ။

၂။ Calculate the number of time units between two dates ကို ရွေးချယ်ပြီး Next ကို နှိပ်ပါ။

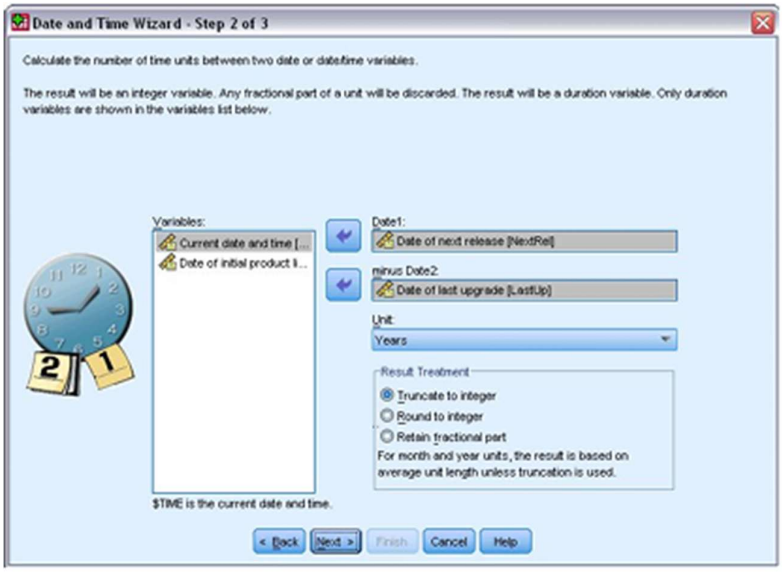

Date and Time Wizard - Step 2 of 3
Calculate the number of time units between two date or datetime variables.
The result will be an integer variable. Any fractional part of a unit will be discarded. The result will be a duration variable. Only duration variables are shown in the variables list below.
Variables:
Current date and time [...
Date of initial product li...
Date1:
Date of next release [NextRel]
minus Date2:
Date of last upgrade [LastUp]
Unit:
Years
Result Treatment
Truncate to integer
Round to integer
Retain fractional part
For month and year units, the result is based on average unit length unless truncation is used.
$TIME is the current date and time.
< Back
Next >
Finish
Cancel
Help

Figure 77. Calculating the length of time between two dates: Step 2

၃။ step 2 မှာ Date1 အတွက် Date of next release ကို ရွေးချယ်ပါ။

၄။ Date2 အတွက် Date of last upgrade ကို ရွေးချယ်ပါ။

၅။ Unit အတွက် Years ကို ရွေးချယ်ပြီး၊ Result Treatment အတွက် Truncate to Integer ကို ရွေးချယ်ပါ။ (ဒါတွေက default ရွေးချယ်မှုတွေပါပဲ)။

၆။ Next ကို နှိပ်ပါ။

၇။ step 3 မှာ result variable ရဲ့ နာမည်အတွက် YearsLastUp လို့ ရိုက်ထည့်ပါ။ Result variables တွေက ရှိပြီးသား variable တစ်ခုရဲ့ နာမည်နဲ့ တူနေလို့ မရပါဘူး။

၈။ result variable ရဲ့ label အတွက် Years since last upgrade လို့ ရိုက်ထည့်ပါ။ Result variables တွေအတွက် variable labels တွေကတော့ ထည့်လည်းရ၊ မထည့်လည်းရ (optional) ပါ။

၉။ Create the variable now ဆိုတဲ့ default selection အတိုင်းပဲ ထားခဲ့ပြီး၊ variable အသစ်ကို ဖန်တီးဖို့ Finish ကို နှိပ်ပါ။

Data Editor ထဲမှာ ပြသနေတဲ့ YearsLastUp ဆိုတဲ့ variable အသစ်ဟာ အဲဒီ dates နှစ်ခုကြားက နှစ်အရေအတွက်ကို integer အနေနဲ့ ပြသနေတာ ဖြစ်ပါတယ်။ နှစ်တစ်နှစ်ရဲ့ အပိုင်းကိန်း (fractional parts) တွေကိုတော့ ဖြတ်ထုတ် (truncated) ထားပါတယ်။

Adding a Duration to a Date

Date တစ်ခုကို ၁၀ ရက် ဒါမှမဟုတ် ၁၂ လ လိုမျိုး durations တွေကို ပေါင်းထည့်တာ ဒါမှမဟုတ် နှုတ်တာတွေ လုပ်နိုင်ပါတယ်။ အရင်အပိုင်းက software ကုမ္ပဏီ ဥပမာကိုပဲ ဆက်သွားရရင်၊ customer တစ်ယောက်စီရဲ့ ပထမဆုံး (initial) tech support contract ပြီးဆုံးမယ့် ရက်စွဲကို သတ်မှတ်ဖို့ စဉ်းစားကြည့်ပါ။ upgrade.sav data file ထဲမှာ contract လုပ်ထားတဲ့ support သက်တမ်း (နှစ်နဲ့) အတွက် variable တစ်ခုနဲ့ ပထမဆုံး ဝယ်ယူခဲ့တဲ့ ရက်စွဲအတွက် variable တစ်ခု ပါဝင်ပါတယ်။ ဒါဆိုရင် ဝယ်ယူခဲ့တဲ့ ရက်စွဲပေါ်မှာ support ပေးမယ့် နှစ်အရေအတွက်ကို ပေါင်းထည့်ခြင်းအားဖြင့် ပထမဆုံး support ပြီးဆုံးမယ့် ရက်စွဲကို သတ်မှတ်နိုင်ပါပြီ။

Date တစ်ခုကို duration တစ်ခု ပေါင်းထည့်ဖို့အတွက်-

၁။ Date and Time Wizard ရဲ့ မိတ်ဆက် မျက်နှာပြင်မှာ Calculate with dates and times ကို ရွေးချယ်ပြီး Next ကို နှိပ်ပါ။

၂။ step 1 မှာ Add or subtract a duration from a date ကို ရွေးချယ်ပြီး Next ကို နှိပ်ပါ။

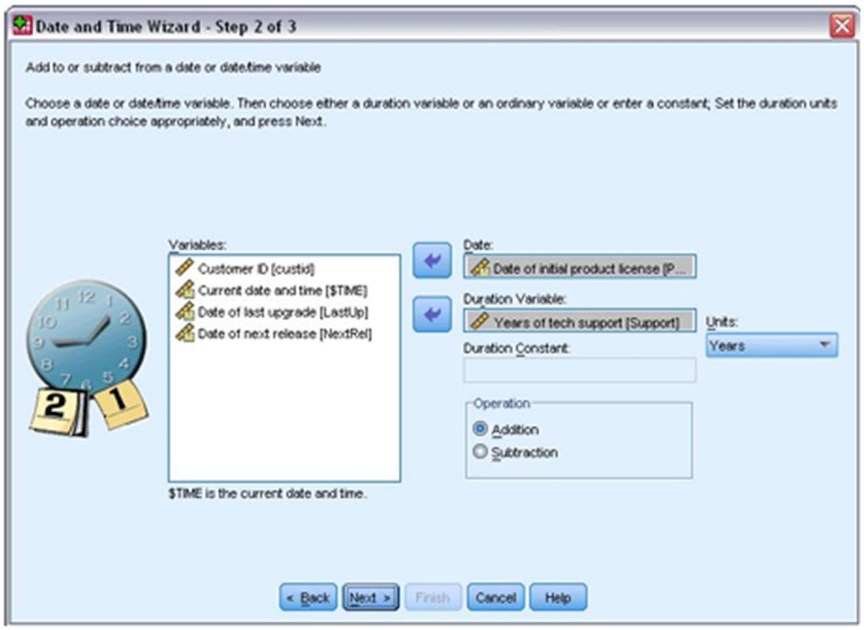

Date and Time Wizard - Step 2 of 3
Add to or subtract from a date or date/time variable
Choose a date or date/time variable. Then choose either a duration variable or an ordinary variable or enter a constant; Set the duration units and operation choice appropriately, and press Next.
Variables:
Customer ID [custid]
Current date and time [$TIME]
Date of last upgrade [LastUp]
Date of next release [NextRel]
Date:
Date of initial product license [P...
Duration Variable:
Years of tech support [Support]
Units:
Years
Duration Constant:
Operation
Addition
Subtraction
$TIME is the current date and time.
< Back
Next >
Finish
Cancel
Help

Figure 78. Adding a duration to a date: Step 2

၃။ Date အတွက် Date of initial product license ကို ရွေးချယ်ပါ။

၄။ step 2 မှာ Duration Variable အတွက် Years of tech support ကို ရွေးချယ်ပါ။ Years of tech support က ရိုးရှင်းတဲ့ numeric variable တစ်ခုသာ ဖြစ်တဲ့အတွက်၊ ဒီ variable ကို duration အနေနဲ့ ပေါင်းထည့်တဲ့အခါ အသုံးပြုမယ့် units တွေကို ညွှန်ပြပေးဖို့ လိုအပ်ပါတယ်။

၅။ Units drop-down list ထဲကနေ Years ကို ရွေးချယ်ပါ။

၆။ Next ကို နှိပ်ပါ။

၇။ step 3 မှာ result variable ရဲ့ နာမည်အတွက် SupEndDate လို့ ရိုက်ထည့်ပါ။ Result variables တွေက ရှိပြီးသား variable တစ်ခုရဲ့ နာမည်နဲ့ တူနေလို့ မရပါဘူး။

၈။ result variable ရဲ့ label အတွက် End date for support လို့ ရိုက်ထည့်ပါ။ Result variables တွေအတွက် variable labels တွေကတော့ ထည့်လည်းရ၊ မထည့်လည်းရ (optional) ပါ။

၉။ Variable အသစ်ကို ဖန်တီးဖို့ Finish ကို နှိပ်ပါ။

Variable အသစ်ကို Data Editor ထဲမှာ ပြသပေးပါလိမ့်မယ်။

Chapter 9. Sorting and Selecting Data

Data files တွေက သင့်ရဲ့ သီးခြားလိုအပ်ချက်တွေအတွက် အမြဲတမ်း အကောင်းဆုံးပုံစံနဲ့ ဖွဲ့စည်းထားတာမျိုး မဟုတ်တတ်ပါဘူး။ Analysis လုပ်ဖို့ data တွေကို ပြင်ဆင်တဲ့အခါ၊ file transformations မျိုးစုံကို ရွေးချယ်လုပ်ဆောင်နိုင်ပါတယ်၊ အဲဒီအထဲမှာ အောက်ပါအချက်တွေ ပါဝင်ပါတယ်-

- Sort data (Data တွေကို အစဉ်လိုက်စီခြင်း) ။ Cases တွေကို variable တစ်ခု ဒါမှမဟုတ် အဲဒီထက်ပိုတဲ့ variables တွေရဲ့ value အပေါ်မူတည်ပြီး အစဉ်လိုက်စီ (sort) နိုင်ပါတယ်။
- Select subsets of cases (Cases အုပ်စုခွဲလေးများကို ရွေးချယ်ခြင်း) ။ သင့်ရဲ့ analysis ကို cases တွေရဲ့ subset တစ်ခုတည်းမှာပဲ ကန့်သတ်လုပ်ဆောင်နိုင်သလို၊ မတူညီတဲ့ subsets တွေအပေါ်မှာ တစ်ပြိုင်နက်တည်း analyses တွေကိုလည်း လုပ်ဆောင်နိုင်ပါတယ်။

ဒီ chapter မှာပါတဲ့ ဥပမာတွေက demo.sav ဆိုတဲ့ data file ကို အသုံးပြုထားတာပါ။ ပိုပြီးသိချင်တယ်ဆိုရင်တော့ Chapter 10 ရဲ့ "Sample Files" ခေါင်းစဉ်မှာ သွားကြည့်နိုင်ပါတယ်။

Sorting Data

Cases တွေကို sort လုပ်တာ (data file ရဲ့ rows တွေကို အစဉ်လိုက်စီတာ) က မကြာခဏဆိုသလို အသုံးဝင်တတ်ပြီး၊ တချို့ analysis အမျိုးအစားတွေအတွက်ဆိုရင် မရှိမဖြစ်လိုအပ်ပါတယ်။ Data file ထဲက cases တွေရဲ့ အစီအစဉ်ကို sorting variables တစ်ခု ဒါမှမဟုတ် အဲဒီထက်ပိုတဲ့ values တွေအပေါ် အခြေခံပြီး ပြန်စီ (reorder) ဖို့အတွက်-

၁။ menus တွေထဲကနေ Data > Sort Cases... ကို ရွေးချယ်ပါ။

Sort Cases dialog box ပေါ်လာပါလိမ့်မယ်။

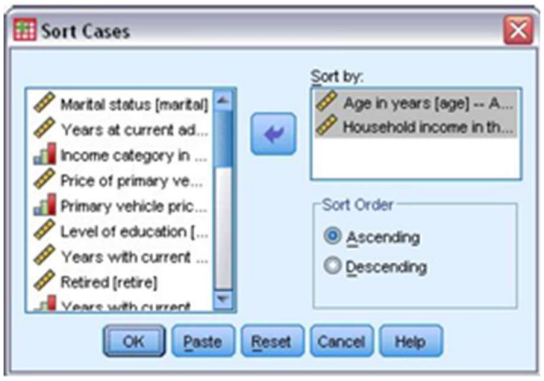


Figure 79. Sort Cases dialog box

၂။ Age in years [age] နဲ့ Household income in thousands [income] variables တွေကို Sort by list ထဲကို ထည့်ပါ။

တကယ်လို့ သင်က sort variables အများကြီးကို ရွေးချယ်ထားမယ်ဆိုရင်၊ Sort by list ထဲမှာ သူတို့ပေါ်နေတဲ့ အစီအစဉ်က cases တွေကို ဘယ်လိုအစီအစဉ်နဲ့ sort လုပ်မလဲဆိုတာကို ဆုံးဖြတ်ပေးပါတယ်။ ဒီဥပမာမှာဆိုရင် Sort by list ထဲက ထည့်သွင်းထားချက်တွေအရ၊ cases တွေကို Age in years [age] ရဲ့ categories တွေအတွင်းမှာ Household income in thousands [income] ရဲ့ value အလိုက် sort လုပ်သွားမှာ ဖြစ်ပါတယ်။ String variables တွေအတွက်ဆိုရင်တော့၊ sort order မှာ အက္ခရာအကြီး (uppercase letters) တွေက အက္ခရာအသေး (lowercase) တွေရဲ့အရှေ့ကနေ လာမှာပါ (ဥပမာအနေနဲ့၊ sort order မှာ string value Yes ဟာ yes ရဲ့အရှေ့ကနေ လာမှာ ဖြစ်ပါတယ်)။

Split-File Processing

Analysis လုပ်ဖို့အတွက် သင့်ရဲ့ data file ကို သီးခြား groups လေးတွေအဖြစ် ခွဲထုတ် (split) ဖို့-

၁။ menus တွေထဲကနေ Data > Split File... ကို ရွေးချယ်ပါ။

Split File dialog box ပေါ်လာပါလိမ့်မယ်။

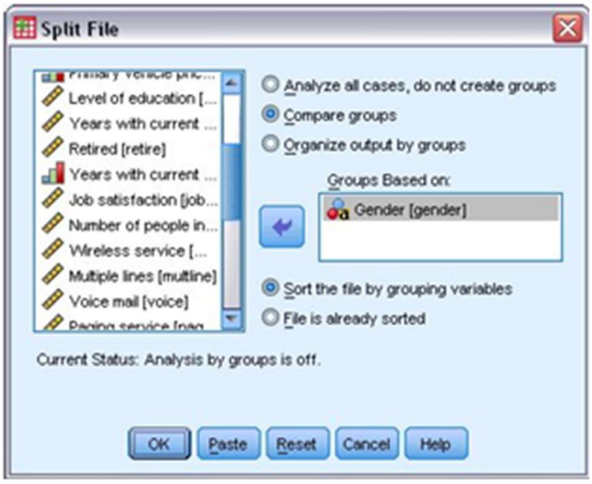


Figure 80. Split File dialog box

၂။ Compare groups ဒါမှမဟုတ် Organize output by groups ကို ရွေးချယ်ပါ။ (ဒီအဆင့်တွေရဲ့နောက်က ဥပမာတွေမှာ အဲဒီ options နှစ်ခုရဲ့ ကွာခြားချက်တွေကို ပြသထားပါတယ်)။

၃။ ဖိုင်ကို ဒီ variables တွေအတွက် သီးခြား groups တွေအဖြစ် ခွဲထုတ်ဖို့ Gender [gender] ကို ရွေးချယ်ပါ။

Grouping variables တွေအနေနဲ့ numeric, short string နဲ့ long string variables တွေကို အသုံးပြုနိုင်ပါတယ်။ Grouping variables တွေကနေ သတ်မှတ်ပေးလိုက်တဲ့ subgroup တစ်ခုစီအတွက် သီးခြား analysis တစ်ခုကို လုပ်ဆောင်ပေးမှာပါ။ တကယ်လို့ သင်က grouping variables အများကြီးကို ရွေးချယ်ထားမယ်ဆိုရင်၊ Groups Based on list ထဲမှာ သူတို့ပေါ်နေတဲ့ အစီအစဉ်က cases တွေကို ဘယ်လိုနည်းလမ်းနဲ့ group ဖွဲ့မလဲဆိုတာကို ဆုံးဖြတ်ပေးမှာ ဖြစ်ပါတယ်။

တကယ်လို့ သင်က Compare groups ကို ရွေးချယ်မယ်ဆိုရင်၊ Frequencies procedure ကနေ ထုတ်ပေးလိုက်တဲ့ summary statistics table မှာ ပြထားသလိုပဲ၊ split-file groups အားလုံးက results တွေကို table (ဒါမှမဟုတ် tables) တစ်ခုတည်းထဲမှာ အတူတူ ထည့်သွင်းပြသပေးမှာ ဖြစ်ပါတယ်။

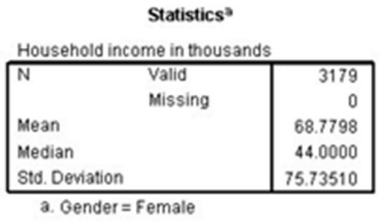
Statistics

Household income in thousands

Female	N	Valid	3179
		Missing	0
	Mean		68.7798
	Median		44.0000
	Std. Deviation		75.73510
Male	N	Valid	3221
		Missing	0
	Mean		70.1608
	Median		45.0000
	Std. Deviation		81.56216

Figure 81. Split-file output with single pivot table

တကယ်လို့ သင်က Organize output by groups ကို ရွေးချယ်ပြီး Frequencies procedure ကို run မယ်ဆိုရင်တော့၊ pivot tables နှစ်ခုကို ဖန်တီးပေးမှာပါ- အမျိုးသမီးများ (females) အတွက် table တစ်ခုနဲ့ အမျိုးသားများ (males) အတွက် table တစ်ခု ဖြစ်ပါတယ်။



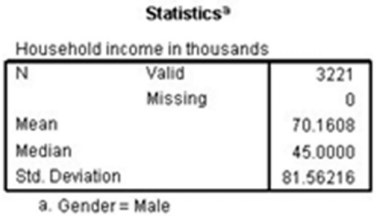
Statistics[a]

Household income in thousands

N	Valid	3179
	Missing	0
Mean		68.7798
Median		44.0000
Std. Deviation		75.73510

a. Gender = Female

Figure 82. Split-file output with pivot table for females



Statistics[a]

Household income in thousands

N	Valid	3221
	Missing	0
Mean		70.1608
Median		45.0000
Std. Deviation		81.56216

a. Gender = Male

Figure 83. Split-file output with pivot table for males

Sorting Cases for Split-File Processing

Split File procedure က grouping variables တွေထဲက တစ်ခုခုအတွက် မတူညီတဲ့ value တစ်ခုကို တွေ့တိုင်း အသစ်တစ်ခုဖြစ်တဲ့ subgroup တစ်ခုကို ဖန်တီးပေးပါတယ်။ ဒါကြောင့်မို့လို့ split-file processing ကို မလုပ်ဆောင်ခင် grouping variables တွေရဲ့ values တွေအပေါ် အခြေခံပြီး cases တွေကို အရင် sort လုပ်ထားဖို့က အရေးကြီးပါတယ်။

ပုံမှန်အားဖြင့် (By default) Split File က data file ကို grouping variables တွေရဲ့ values တွေအပေါ် အခြေခံပြီး အလိုအလျောက် sort လုပ်ပေးပါတယ်။ တကယ်လို့ ဖိုင်က မှန်ကန်တဲ့ အစီအစဉ်နဲ့ sort လုပ်ပြီးသား ဖြစ်နေမယ်ဆိုရင်၊ File is already sorted ကို ရွေးချယ်ခြင်းအားဖြင့် processing time ကို သက်သာစေနိုင်ပါတယ်။

Turning Split-File Processing On and Off

Split-file processing ကို စတင်လုပ်ဆောင်လိုက်ပြီးရင်၊ သင်က သူ့ကို ပြန်မပိတ် (turn off) မချင်း အဲဒီ session တစ်ခုလုံးအတွက် ဆက်လက် အသက်ဝင်နေမှာ ဖြစ်ပါတယ်။

- Analyze all cases. ဒီ option က split-file processing ကို ပိတ် (turn off) ပေးပါတယ်။

- Compare groups နဲ့ Organize output by groups. ဒီ option ကတော့ split-file processing ကို ဖွင့် (turn on) ပေးပါတယ်။

တကယ်လို့ split-file processing က အသက်ဝင်နေမယ်ဆိုရင်၊ application window ရဲ့ အောက်ခြေက status bar မှာ Split File on ဆိုတဲ့ message လေး ပေါ်နေမှာ ဖြစ်ပါတယ်။

Selecting Subsets of Cases

Variables တွေနဲ့ ရှုပ်ထွေးတဲ့ expressions တွေ ပါဝင်တဲ့ စံသတ်မှတ်ချက် (criteria) တွေအပေါ် အခြေခံပြီး သင့်ရဲ့ analysis ကို သီးခြား subgroup တစ်ခုတည်းမှာပဲ ကန့်သတ်ထားနိုင်ပါတယ်။ Random sample (ကျပန်းရွေးချယ်ထားတဲ့ cases များ) ကိုလည်း ရွေးချယ်နိုင်ပါသေးတယ်။ Subgroup တစ်ခုကို သတ်မှတ်ဖို့ အသုံးပြုတဲ့ criteria တွေထဲမှာ အောက်ပါတို့ ပါဝင်နိုင်ပါတယ်-

- Variable values တွေနဲ့ ranges တွေ
- Date နဲ့ time ranges တွေ
- Case (row) numbers တွေ
- Arithmetic expressions တွေ
- Logical expressions တွေ
- Functions တွေ

Analysis အတွက် cases တွေရဲ့ subset တစ်ခုကို ရွေးချယ်ဖို့-

၁။ menus တွေထဲကနေ Data > Select Cases... ကို ရွေးချယ်ပါ။

ဒါက Select Cases dialog box ကို ဖွင့်ပေးပါလိမ့်မယ်။

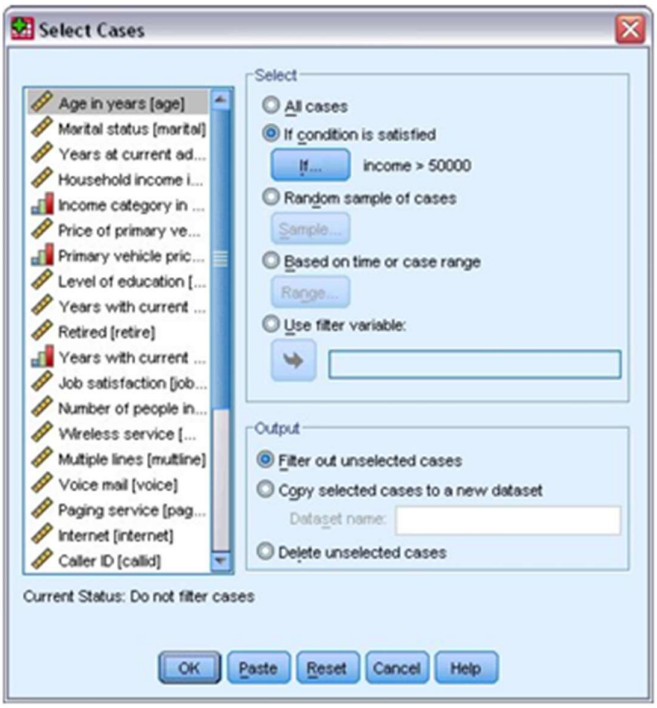


Figure 84. Select Cases dialog box

Selecting Cases Based on Conditional Expressions

Conditional expression တစ်ခုပေါ် အခြေခံပြီး cases တွေကို ရွေးချယ်ဖို့အတွက်-

၁။ Select Cases dialog box ထဲမှာ If condition is satisfied ကို ရွေးချယ်ပြီး If ကို နှိပ်ပါ။ ဒါက Select Cases If dialog box ကို ဖွင့်ပေးပါလိမ့်မယ်။

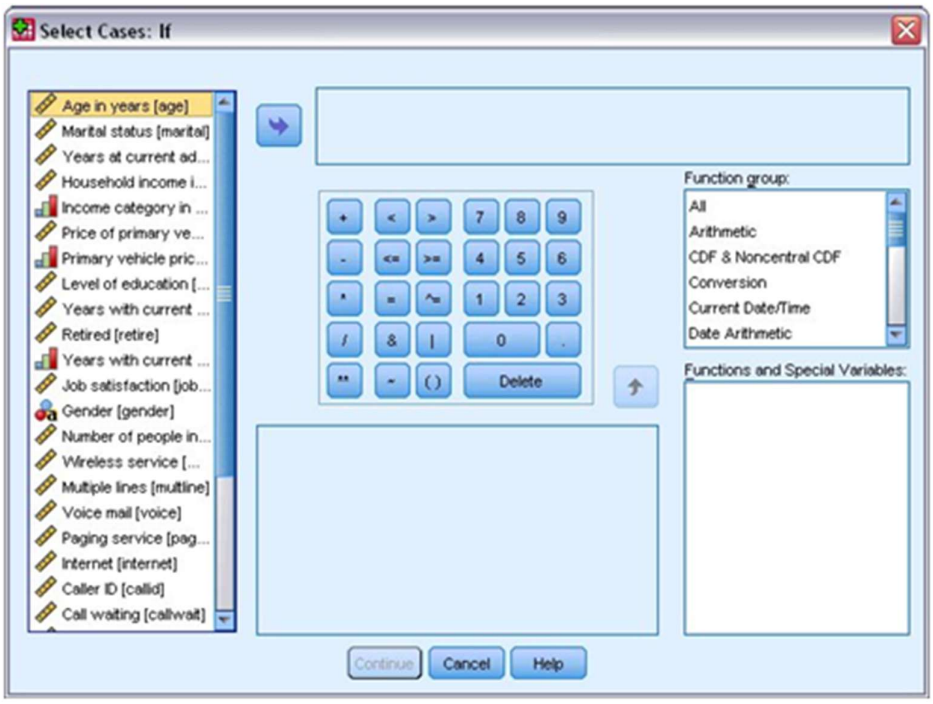


Figure 85. Select Cases If dialog box

Conditional expression ထဲမှာ ရှိပြီးသား variable names တွေ၊ constants တွေ၊ arithmetic operators တွေ၊ logical operators တွေ၊ relational operators တွေနဲ့ functions တွေကို အသုံးပြုနိုင်ပါတယ်။ Output window ထဲက text တွေလိုပဲ text box ထဲမှာ expression ကို ရိုက်ထည့်တာ၊ edit လုပ်တာတွေ လုပ်နိုင်ပါတယ်။ Expression ထဲကို elements တွေ paste လုပ်ထည့်ဖို့အတွက် calculator pad, variable list နဲ့ function list တွေကိုလည်း အသုံးပြုနိုင်ပါတယ်။ ပိုပြီးသိချင်တယ်ဆိုရင် "Using Conditional Expressions" ဆိုတဲ့ ခေါင်းစဉ်မှာ သွားကြည့်နိုင်ပါတယ်။

Selecting a Random Sample

Random sample တစ်ခု ရယူဖို့အတွက်-

၁။ Select Cases dialog box ထဲမှာ Random sample of cases ကို ရွေးချယ်ပါ။

၂။ Sample ကို နှိပ်ပါ။

ဒါက Select Cases Random Sample dialog box ကို ဖွင့်ပေးပါလိမ့်မယ်။

Figure 86. Select Cases Random Sample dialog box

Sample size အတွက် အောက်ပါ ရွေးချယ်စရာတွေထဲက တစ်ခုကို ရွေးချယ်နိုင်ပါတယ်-

- **Approximately (ခန့်မှန်းခြေ)။** User သတ်မှတ်ပေးထားတဲ့ ရာခိုင်နှုန်း (percentage) တစ်ခုပါ။ ဒီ option က သတ်မှတ်ထားတဲ့ cases ရာခိုင်နှုန်းရဲ့ ခန့်မှန်းခြေ random sample တစ်ခုကို ထုတ်ပေးပါတယ်။
- **Exactly (တိကျစွာ)။** User သတ်မှတ်ပေးထားတဲ့ cases အရေအတွက်ပါ။ Sample ထုတ်ယူမယ့် cases အရေအတွက်ကိုလည်း သတ်မှတ်ပေးဖို့ လိုအပ်ပါတယ်။ ဒီ ဒုတိယ ကိန်းဂဏန်းက data file ထဲမှာရှိတဲ့ စုစုပေါင်း cases အရေအတွက်ထက် နည်းနေဖို့ ဒါမှမဟုတ် ညီမျှနေဖို့ လိုအပ်ပါတယ်။ တကယ်လို့ အဲဒီကိန်းဂဏန်းက data file ထဲက စုစုပေါင်း cases အရေအတွက်ထက် ကျော်လွန်နေမယ်ဆိုရင်၊ sample ထဲမှာ သင်တောင်းဆိုထားတဲ့ အရေအတွက်ထက် အချိုးကျ ပိုနည်းတဲ့ cases တွေပဲ ပါဝင်လာမှာ ဖြစ်ပါတယ်။

Selecting a Time Range or Case Range

Dates တွေ၊ times တွေ ဒါမှမဟုတ် observation (row) numbers တွေအပေါ် အခြေခံပြီး cases တွေရဲ့ range တစ်ခုကို ရွေးချယ်ဖို့-

၁။ Select Cases dialog box ထဲမှာ Based on time or case range ကို ရွေးချယ်ပြီး Range ကို နှိပ်ပါ။ ဒါက Select Cases Range dialog box ကို ဖွင့်ပေးမှာဖြစ်ပြီး၊ အဲဒီအထဲမှာ observation (row) numbers တွေရဲ့ range တစ်ခုကို ရွေးချယ်နိုင်ပါတယ်။

Figure 87. Select Cases Range dialog box

- **First Case (ပထမဆုံး Case) ။** Range အတွက် စတင်မယ့် date ဒါမှမဟုတ် time values တွေကို ရိုက်ထည့်ပါ။ တကယ်လို့ date variables တွေကို သတ်မှတ်မထားဘူးဆိုရင်၊ စတင်မယ့် observation number (Split File ဖွင့်မထားဘူးဆိုရင် Data Editor ထဲက row number ပါ) ကို ရိုက်ထည့်ပါ။ တကယ်လို့ သင်က Last Case value ကို သတ်မှတ်မပေးထားဘူးဆိုရင်၊ စတင်တဲ့ date/time ကနေစပြီး time series ရဲ့ အဆုံးအထိ cases အားလုံးကို ရွေးချယ်သွားမှာပါ။
- **Last Case (နောက်ဆုံး Case) ။** Range အတွက် အဆုံးသတ်မယ့် date ဒါမှမဟုတ် time values တွေကို ရိုက်ထည့်ပါ။ တကယ်လို့ date variables တွေကို သတ်မှတ်မထားဘူးဆိုရင်၊ အဆုံးသတ်မယ့် observation number (Split File ဖွင့်မထားဘူးဆိုရင် Data Editor ထဲက row number ပါ) ကို ရိုက်ထည့်ပါ။
- တကယ်လို့ သင်က First Case value ကို သတ်မှတ်မပေးထားဘူးဆိုရင်၊ time series ရဲ့ အစကနေ အဆုံးသတ်မယ့် date/time အထိ cases အားလုံးကို ရွေးချယ်သွားမှာပါ။

Date variables တွေကို သတ်မှတ်ပေးထားတဲ့ time series data အတွက်ဆိုရင်၊ အဲဒီသတ်မှတ်ထားတဲ့ date variables တွေအပေါ် အခြေခံပြီး dates ဒါမှမဟုတ် times တွေရဲ့ range တစ်ခုကို ရွေးချယ်နိုင်ပါတယ်။

Figure 88. Select Cases Range dialog box (time series)

Case တစ်ခုစီက မတူညီတဲ့ အချိန်တစ်ခုက observations တွေကို ကိုယ်စားပြုပြီး၊ ဖိုင်ကိုလည်း အချိန်အစဉ်လိုက် (chronological order) အတိုင်း sort လုပ်ထားပါတယ်။

Time series data အတွက် date variables တွေ ဖန်တီး (generate) ဖို့-

၂။ menus တွေထဲကနေ Data > Define Dates... ကို ရွေးချယ်ပါ။

Treatment of Unselected Cases

ရွေးချယ်မခံရတဲ့ (unselected) cases တွေကို ဘယ်လိုကိုင်တွယ်ဖြေရှင်းမလဲ ဆိုတာနဲ့ပတ်သက်ပြီး အောက်ပါ ရွေးချယ်စရာတွေထဲက တစ်ခုကို ရွေးချယ်နိုင်ပါတယ်-

- **Filter out unselected cases (ရွေးချယ်မခံရတဲ့ cases တွေကို စစ်ထုတ်ခြင်း)။** Unselected cases တွေဟာ analysis ထဲမှာ ပါဝင်မှာ မဟုတ်ပေမယ့် dataset ထဲမှာတော့ ကျန်ရှိနေဦးမှာပါ။ တကယ်လို့ သင်က filtering ကို ပိတ် (turn off) လိုက်မယ်ဆိုရင် ဒီ session ရဲ့ နောက်ပိုင်းမှာ အဲဒီ unselected cases တွေကို ပြန်သုံးနိုင်ပါတယ်။ တကယ်လို့ သင်က random sample တစ်ခုကို ရွေးချယ်တာ ဒါမှမဟုတ် conditional expression တစ်ခုအပေါ် အခြေခံပြီး cases တွေကို ရွေးချယ်မယ်ဆိုရင်၊ filter_$ လို့ နာမည်ပေးထားတဲ့ variable တစ်ခုကို ဖန်တီးပေးမှာဖြစ်ပြီး၊ selected cases တွေအတွက် value 1 နဲ့ unselected cases တွေအတွက် value 0 ဆိုပြီး သတ်မှတ်ပေးမှာ ဖြစ်ပါတယ်။

- **Copy selected cases to a new dataset (ရွေးချယ်ထားတဲ့ cases တွေကို dataset အသစ်ဆီ ကူးယူခြင်း) ။** Selected cases တွေကို dataset အသစ်တစ်ခုဆီ copy ကူးထည့်ပေးမှာဖြစ်ပြီး၊ မူရင်း dataset ကိုတော့ ဘာသက်ရောက်မှုမှ မရှိစေပါဘူး။ Unselected cases တွေကတော့ dataset အသစ်ထဲမှာ ပါဝင်မှာ မဟုတ်ဘဲ မူရင်း dataset ထဲမှာပဲ သူတို့ရဲ့ မူလအခြေအနေအတိုင်း ကျန်ရှိနေမှာ ဖြစ်ပါတယ်။

- **Delete unselected cases (ရွေးချယ်မခံရတဲ့ cases တွေကို ဖျက်ပစ်ခြင်း)။** Unselected cases တွေကို dataset ထဲကနေ delete လုပ်ပစ်မှာပါ။ ဖျက်လိုက်တဲ့ cases တွေကို ပြန်လိုချင်တယ်ဆိုရင်၊ ဘာအပြောင်းအလဲကိုမှ save မလုပ်ဘဲ ဖိုင်ထဲကနေ ထွက်ပြီး ဖိုင်ကို ပြန်ဖွင့်မှပဲ ပြန်ရနိုင်ပါမယ်။ တကယ်လို့ သင်က data file ထဲက အပြောင်းအလဲတွေကို save လုပ်လိုက်မယ်ဆိုရင်တော့ cases တွေကို အပြီးတိုင် ဖျက်ပစ်လိုက်မှာ ဖြစ်ပါတယ်။ မှတ်ချက် - တကယ်လို့ သင်က unselected cases တွေကို ဖျက်ပစ်ပြီး ဖိုင်ကို save လိုက်မယ်ဆိုရင်၊ အဲဒီ cases တွေကို ပြန်လည်ရယူ (recover) နိုင်တော့မှာ မဟုတ်ပါဘူး။

Case Selection Status

တကယ်လို့ သင်က cases တွေရဲ့ subset တစ်ခုကို ရွေးချယ်ထားပေမယ့် unselected cases တွေကို မစွန့်ပစ် (discard) ဘူးဆိုရင်၊ Data Editor ထဲမှာ unselected cases တွေရဲ့ row number ပေါ်မှာ ဒေါင်လိုက်မျဉ်းစောင်းလေး (diagonal line) တားပြီး မှတ်သားပြသပေးထားမှာ ဖြစ်ပါတယ်။

Figure 89. Case selection status

Chapter 10. Sample Files (နမူနာ ဖိုင်များ)

ဒီ Chapter 10 မှာတော့ Sample Files တွေအကြောင်းကို ရှင်းပြပေးသွားမှာပါ။

Product နဲ့အတူ install လုပ်ထားတဲ့ sample files တွေကို installation directory ရဲ့ Samples subdirectory ထဲမှာ ရှာတွေ့နိုင်ပါတယ်။ အင်္ဂလိပ်၊ ပြင်သစ်၊ ဂျာမန်၊ အီတလီ၊ ဂျပန်၊ ကိုရီးယား၊ ပိုလန်၊ ရုရှား၊ ရိုးရှင်းသောတရုတ်၊ စပိန်နဲ့ ရိုးရာတရုတ်ဘာသာစကား တစ်ခုချင်းစီအတွက် Samples subdirectory ထဲမှာ သီးခြား folder တစ်ခုစီ ရှိပါတယ်။ Sample files အားလုံးကို ဘာသာစကားတိုင်းအတွက် မရနိုင်ပါဘူး။ တကယ်လို့ ဘာသာစကားတစ်ခုမှာ sample file တစ်ခု မရှိဘူးဆိုရင်၊ အဲဒီ ဘာသာစကား folder ထဲမှာ အင်္ဂလိပ် version နဲ့ sample file ကိုပဲ ထည့်သွင်းပေးထားပါတယ်။

Descriptions (ဖော်ပြချက်များ)

အောက်ပါတို့ကတော့ documentation တစ်လျှောက်က ဥပမာအမျိုးမျိုးမှာ အသုံးပြုထားတဲ့ sample files တွေရဲ့ အကျဉ်းချုပ် ရှင်းလင်းချက်တွေပဲ ဖြစ်ပါတယ်။

- **accidents.sav.** ဒါက သတ်မှတ်ထားတဲ့ ဒေသတစ်ခုအတွင်းမှာရှိတဲ့ ကားအက်ဆီးဒင့်တွေအတွက် age နဲ့ gender risk factors တွေကို လေ့လာနေတဲ့ အာမခံကုမ္ပဏီတစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီဟာ age category နဲ့ gender တို့ရဲ့ cross-classification ကို ကိုယ်စားပြုပါတယ်။

- **adl.sav.** ဒါက လေဖြတ်လူနာတွေ (stroke patients) အတွက် အဆိုပြုထားတဲ့ therapy အမျိုးအစားတစ်ခုရဲ့ အကျိုးကျေးဇူးတွေကို ဆုံးဖြတ်ဖို့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ ဆရာဝန်တွေက အမျိုးသမီး လေဖြတ်လူနာတွေကို အုပ်စုနှစ်ခုထဲက တစ်ခုကို random အနေနဲ့ သတ်မှတ်ပေးခဲ့ပါတယ်။ ပထမအုပ်စုက standard physical therapy ကို ရရှိပြီး၊ ဒုတိယအုပ်စုကတော့ ထပ်ဆောင်း emotional therapy ကို ရရှိပါတယ်။ Treatments တွေပေးပြီး သုံးလအကြာမှာ၊ လူနာတစ်ယောက်ချင်းစီရဲ့ နေ့စဉ်ဘဝ လုပ်ငန်းဆောင်တာတွေ လုပ်နိုင်စွမ်းကို ordinal variables တွေအနေနဲ့ score မှတ်သားထားပါတယ်။

- **advert.sav.** ဒါက ကြော်ငြာအတွက် သုံးစွဲခဲ့တဲ့ ငွေကြေးနဲ့ အဲဒီကနေ ရရှိလာတဲ့ အရောင်း (sales) ကြားက ဆက်စပ်မှုကို စစ်ဆေးဖို့ လက်လီရောင်းချသူ (retailer) တစ်ဦးရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ ဒီအတွက် သူတို့က အတိတ်က sales figures တွေနဲ့ ဆက်စပ်နေတဲ့ advertising costs တွေကို ကောက်ယူထားပါတယ်။

- **aflatoxin.sav.** ဒါက ပြောင်းဖူးအထွက်နှုန်း (corn crops) တွေမှာ အဆိပ်တစ်မျိုးဖြစ်တဲ့ aflatoxin ပါဝင်မှုကို စမ်းသပ်ခြင်းနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုဖြစ်ပြီး၊ အဲဒီအဆိပ်ရဲ့ concentration ဟာ crop yields တွေကြားမှာရော၊ အတွင်းမှာပါ အလွန်ကွာခြားမှုရှိပါတယ်။ Grain processor တစ်ဦးဟာ crop yields ၈ ခုစီကနေ samples ၁၆ ခုစီကို လက်ခံရရှိပြီး၊ alfatoxin levels တွေကို parts per billion (PPB) နဲ့ တိုင်းတာထားပါတယ်။

- **anorectic.sav.** Anorectic/bulimic behavior ရဲ့ standardized symptomatology ကို လုပ်ဆောင်နေစဉ်အတွင်း၊ သုတေသီတွေက စားသောက်မှုပုံစံ မမှန်တဲ့ (eating disorders) ဆယ်ကျော်သက် ၅၅ ဦးကို လေ့လာခဲ့ပါတယ်။ လူနာတစ်ဦးစီကို ၄ နှစ်အတွင်းမှာ ၄ ကြိမ် တွေ့ဆုံခဲ့ပြီး စုစုပေါင်း 220 observations ရရှိခဲ့ပါတယ်။ Observation တစ်ခုစီမှာ လူနာတွေကို symptoms ၁၆ ခုစီအတွက် score သတ်မှတ်ပေးခဲ့ပါတယ်။ Patient 71 ရဲ့ time 2 မှာ၊ patient 76 ရဲ့ time 2 မှာနဲ့ patient 47 ရဲ့ time 3 မှာ symptom scores တွေ missing ဖြစ်နေတဲ့အတွက် 217 valid observations သာ ကျန်ရှိပါတယ်။

- **bankloan.sav.** ဒါက loan defaults (ချေးငွေ ပြန်မဆပ်နိုင်မှု) နှုန်းကို လျှော့ချဖို့ ဘဏ်တစ်ခုရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ ဒီဖိုင်ထဲမှာ အရင်က customers တွေနဲ့ အလားအလာရှိတဲ့ customers ၈၅၀ ရဲ့ financial နဲ့ demographic သတင်းအချက်အလက်တွေ ပါဝင်ပါတယ်။ ပထမ cases ၇၀၀ က ယခင်က ချေးငွေရယူခဲ့တဲ့ customers တွေပါ။ နောက်ဆုံး cases ၁၅၀ ကတော့ ဘဏ်အနေနဲ့ good ဒါမှမဟုတ် bad credit risks လို့ classify လုပ်ဖို့လိုအပ်နေတဲ့ အလားအလာရှိတဲ့ customers တွေ ဖြစ်ပါတယ်။

- **bankloan_binning.sav.** ဒါက ယခင် customers ၅၀၀၀ ရဲ့ financial နဲ့ demographic သတင်းအချက်အလက်တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။

- **behavior.sav.** Classic example တစ်ခုမှာ ကျောင်းသား ၅၂ ဦးကို အခြေအနေ (situations) ၁၅ မျိုးနဲ့ အပြုအမူ (behaviors) ၁၅ မျိုးတို့ရဲ့ ပေါင်းစပ်မှုတွေကို 0="extremely appropriate" ကနေ 9="extremely inappropriate" ဆိုတဲ့ 10-point scale ပေါ်မှာ rate လုပ်ဖို့ မေးမြန်းခဲ့ပါတယ်။ တစ်ဦးချင်းစီအတွက် ပျမ်းမျှယူလိုက်တဲ့ (Averaged over individuals) values တွေကို dissimilarities တွေအနေနဲ့ ယူဆပါတယ်။

- **behavior_ini.sav.** ဒီ data file ထဲမှာ behavior.sav အတွက် two-dimensional solution တစ်ခုရဲ့ initial configuration ပါဝင်ပါတယ်။

- **brakes.sav.** ဒါက စွမ်းဆောင်ရည်မြင့် ကားတွေအတွက် disc brakes တွေ ထုတ်လုပ်တဲ့ စက်ရုံတစ်ရုံရဲ့ quality control နဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Data file ထဲမှာ production machines ၈ လုံးစီကနေ ထုတ်လုပ်တဲ့ discs ၁၆ ခုစီရဲ့ အချင်း (diameter) တိုင်းတာချက်တွေ ပါဝင်ပါတယ်။ Brakes တွေအတွက် target diameter က ၃၂၂ မီလီမီတာ ဖြစ်ပါတယ်။

- **breakfast.sav.** Classic study တစ်ခုမှာ Wharton School က MBA ကျောင်းသား ၂၁ ဦးနဲ့ သူတို့ရဲ့ အိမ်ထောင်ဖက်တွေကို breakfast items ၁၅ ခုကို 1="most preferred" ကနေ 15="least preferred" အထိ preference အလိုက် rank လုပ်ခိုင်းခဲ့ပါတယ်။ သူတို့ရဲ့ preferences တွေကို "Overall preference" ကနေ "Snack, with beverage only" အထိ မတူညီတဲ့ scenarios ၆ ခုအောက်မှာ မှတ်တမ်းတင်ခဲ့ပါတယ်။

- **breakfast-overall.sav.** ဒီ data file ထဲမှာ ပထမဆုံး scenario ဖြစ်တဲ့ "Overall preference" တစ်ခုတည်းအတွက် breakfast item preferences တွေသာ ပါဝင်ပါတယ်။

- **broadband_1.sav.** ဒါက နိုင်ငံလုံးဆိုင်ရာ broadband service တစ်ခုအတွက် ဒေသအလိုက် subscribers အရေအတွက်တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။ Data file ထဲမှာ ၄ နှစ်တာကာလအတွင်း ဒေသ ၈၅ ခုအတွက် လစဉ် subscriber numbers တွေ ပါဝင်ပါတယ်။

- **broadband_2.sav.** ဒီ data file က broadband_1.sav နဲ့ အတူတူပဲဖြစ်ပေမယ့် နောက်ထပ် ၃ လအတွက် data တွေ ထပ်ပေါင်းပါဝင်ပါတယ်။

- **car_insurance_claims.sav.** ကားတွေအတွက် damage claims တွေနဲ့ သက်ဆိုင်တဲ့ dataset တစ်ခုဖြစ်ပါတယ်။ ပျမ်းမျှ claim amount ကို policyholder age, vehicle type နဲ့ vehicle age တို့ရဲ့ linear combination ကို dependent variable ရဲ့ mean နဲ့ ဆက်စပ်ပေးဖို့ inverse link function ကိုသုံးပြီး gamma distribution ရှိတယ်လို့ model လုပ်လို့ရပါတယ်။ တောင်းခံလာတဲ့ claims အရေအတွက်ကို scaling weight အနေနဲ့ သုံးနိုင်ပါတယ်။

- **car_sales.sav.** ဒီ data file ထဲမှာ ယာဉ်အမျိုးအစားနဲ့ မော်ဒယ်အမျိုးမျိုးအတွက် hypothetical sales estimates တွေ၊ list prices တွေနဲ့ physical specifications တွေ ပါဝင်ပါတယ်။ List prices နဲ့ physical specifications တွေကို edmunds.com နဲ့ manufacturer sites တွေကနေ တစ်လှည့်စီ ရယူခဲ့တာပါ။

- **car_sales_uprepared.sav.** ဒါက fields တွေရဲ့ transformed versions တွေ ဘာတစ်ခုမှ မပါဝင်တဲ့ car_sales.sav ရဲ့ modified version တစ်ခုပါ။

- **carpet.sav.** လူသိများတဲ့ ဥပမာတစ်ခုမှာ ကော်ဇောသန့်စင်ဆေးရည် အသစ်တစ်ခုကို ဈေးကွက်တင်ဖို့ စိတ်ဝင်စားနေတဲ့ ကုမ္ပဏီတစ်ခုဟာ consumer preference အပေါ် package design, brand name, price, Good Housekeeping seal နဲ့ money-back guarantee ဆိုတဲ့ factors ၅ ခုရဲ့ လွှမ်းမိုးမှုကို လေ့လာချင်ပါတယ်။ Package design အတွက် applicator brush နေရာ မတူညီတဲ့ factor levels ၃ ခု၊ brand names ၃ ခု (K2R, Glory, နဲ့ Bissell)၊ price levels ၃ ခု နဲ့ နောက်ဆုံး factors ၂ ခုအတွက် (no ဒါမှမဟုတ် yes ဆိုတဲ့) levels ၂ ခုစီ ရှိပါတယ်။ စားသုံးသူ ၁၀ ဦးက ဒီ factors တွေနဲ့ သတ်မှတ်ထားတဲ့ profiles ၂၂ ခုကို rank လုပ်ခဲ့ကြပါတယ်။ Preference ဆိုတဲ့ variable ထဲမှာ profile တစ်ခုစီအတွက် average rankings ရဲ့ rank ပါဝင်ပါတယ်။ Low rankings တွေဟာ high preference ကို ကိုယ်စားပြုပါတယ်။ ဒီ variable ဟာ profile တစ်ခုချင်းစီအတွက် overall measure of preference ကို ထင်ဟပ်စေပါတယ်။

- **carpet_prefs.sav.** ဒီ data file က carpet.sav မှာ ရှင်းပြထားတဲ့ ဥပမာအတိုင်း အခြေခံထားတာဖြစ်ပေမယ့်၊ စားသုံးသူ ၁၀ ဦးစီဆီကနေ ကောက်ယူထားတဲ့ အမှန်တကယ် rankings တွေ ပါဝင်ပါတယ်။ စားသုံးသူတွေကို product profiles ၂၂ ခုကို အကြိုက်ဆုံးကနေ အနည်းဆုံးအထိ rank လုပ်ခိုင်းခဲ့တာပါ။ Variables PREF1 ကနေ PREF22 အထိက carpet_plan.sav မှာ သတ်မှတ်ထားတဲ့အတိုင်း သက်ဆိုင်ရာ profiles တွေရဲ့ identifiers တွေ ပါဝင်ပါတယ်။

- **catalog.sav.** ဒီ data file ထဲမှာ catalog ကုမ္ပဏီတစ်ခုကနေ ရောင်းချတဲ့ products ၃ မျိုးအတွက် hypothetical monthly sales figures တွေ ပါဝင်ပါတယ်။ ဖြစ်နိုင်ခြေရှိတဲ့ predictor variables ၅ ခုအတွက် data တွေလည်း ပါဝင်ပါတယ်။

- **catalog_seasfac.sav.** ဒီ data file က catalog.sav နဲ့ အတူတူပဲဖြစ်ပေမယ့် Seasonal Decomposition procedure ကနေ တွက်ချက်ထားတဲ့ seasonal factors တွေနဲ့အတူ သက်ဆိုင်ရာ date variables တွေကိုပါ ထည့်သွင်းထားပါတယ်။

- **cellular.sav.** ဒါက churn (ဖောက်သည်လျော့ကျမှု) ကို လျှော့ချဖို့ cellular phone ကုမ္ပဏီတစ်ခုရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Accounts တွေအပေါ် churn propensity scores တွေကို 0 ကနေ 100 အထိ

သတ်မှတ်ပေးထားပါတယ်။ 50 နဲ့ အထက် score ရှိတဲ့ accounts တွေက providers ပြောင်းဖို့ ကြည့်နေနိုင်ပါတယ်။

- **ceramics.sav.** ဒါက premium alloy အသစ်တစ်ခုဟာ standard alloy ထက် ပိုပြီး heat resistance ရှိသလားဆိုတာကို ဆုံးဖြတ်ဖို့ ထုတ်လုပ်သူတစ်ဦးရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက alloys တွေထဲက တစ်ခုကို သီးခြားစမ်းသပ်မှုတစ်ခုကို ကိုယ်စားပြုပြီး၊ bearing ပျက်စီးသွားတဲ့ (failed) အပူချိန်ကို မှတ်တမ်းတင်ထားပါတယ်။

- **cereal.sav.** ဒါက လူ ၈၈၀ ကို သူတို့ရဲ့ breakfast preferences တွေအကြောင်း မေးမြန်းထားတဲ့ poll တစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုဖြစ်ပြီး၊ သူတို့ရဲ့ age, gender, marital status နဲ့ (တစ်ပတ်ကို အနည်းဆုံး နှစ်ကြိမ် လေ့ကျင့်ခန်း လုပ်မလုပ်ပေါ် မူတည်ပြီး) active lifestyle ရှိမရှိကိုပါ မှတ်သားထားပါတယ်။ Case တစ်ခုစီက ဖြေဆိုသူ (respondent) တစ်ဦးစီကို ကိုယ်စားပြုပါတယ်။

- **clothing_defects.sav.** ဒါက အထည်ချုပ်စက်ရုံတစ်ခုက quality control process နဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ စက်ရုံကနေ ထုတ်လုပ်လိုက်တဲ့ lot တစ်ခုစီကနေ inspectors တွေက အထည် samples တွေကို ယူပြီး လက်ခံနိုင်စွမ်းမရှိတဲ့ (unacceptable) အထည်အရေအတွက်ကို ရေတွက်ပါတယ်။

- **coffee.sav.** ဒီ data file က iced-coffee brands ၆ ခုရဲ့ perceived images တွေနဲ့ သက်ဆိုင်ပါတယ်။ Iced-coffee image attributes ၂၃ ခုစီအတွက် လူတွေကို အဲဒီ attribute နဲ့ ကိုက်ညီတဲ့ brands အားလုံးကို ရွေးချယ်ခိုင်းခဲ့ပါတယ်။ Confidentiality ကို ထိန်းသိမ်းထားဖို့ brands ၆ ခုကို AA, BB, CC, DD, EE, နဲ့ FF ဆိုပြီး နာမည်ပေးထားပါတယ်။

- **contacts.sav.** ဒါက corporate computer sales representatives အဖွဲ့တစ်ဖွဲ့အတွက် contact lists တွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Contact တစ်ခုစီကို သူတို့အလုပ်လုပ်တဲ့ ကုမ္ပဏီရဲ့ department နဲ့ သူတို့ရဲ့ company ranks တွေအလိုက် ခွဲခြားသတ်မှတ် (categorized) ထားပါတယ်။ နောက်ဆုံး ရောင်းချခဲ့ရတဲ့ (last sale) ပမာဏ၊ နောက်ဆုံး ရောင်းချပြီးကတည်းက ကြာချိန်နဲ့ contact ရဲ့ ကုမ္ပဏီ အရွယ်အစားတို့ကိုလည်း မှတ်တမ်းတင်ထားပါတယ်။

- **creditpromo.sav.** ဒါက မကြာသေးမီက ပြုလုပ်ခဲ့တဲ့ credit card promotion တစ်ခုရဲ့ ထိရောက်မှုကို အကဲဖြတ်ဖို့ department store တစ်ခုရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ ဒီအတွက် cardholders ၅၀၀ ကို random အနေနဲ့

ရွေးချယ်ခဲ့ပါတယ်။ တစ်ဝက်က နောက် ၃ လအတွင်း ဝယ်ယူမှုတွေအပေါ် အတိုးနှုန်း လျှော့ချပေးမယ့် ad တစ်ခုကို ရရှိခဲ့ပါတယ်။ ကျန်တစ်ဝက်ကတော့ ပုံမှန် seasonal ad ကို ရရှိခဲ့ပါတယ်။

- **customer_dbase.sav.** ဒါက data warehouse ထဲမှာရှိတဲ့ သတင်းအချက်အလက်တွေကို အသုံးပြုပြီး တုံ့ပြန်ဖို့ အလားအလာအများဆုံးရှိတဲ့ customers တွေကို special offers တွေပေးဖို့ ကုမ္ပဏီတစ်ခုရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Customer base ထဲက subset တစ်ခုကို random အနေနဲ့ ရွေးချယ်ပြီး special offers တွေ ပေးခဲ့ကာ၊ သူတို့ရဲ့ တုံ့ပြန်မှု (responses) တွေကို မှတ်တမ်းတင်ထားပါတယ်။
- **customer_information.sav.** အမည်နဲ့ လိပ်စာလိုမျိုး customer mailing သတင်းအချက်အလက်တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။
- **customer_subset.sav.** customer_dbase.sav ထဲက cases ၈၀ ပါဝင်တဲ့ subset တစ်ခုပါ။
- **debate.sav.** ဒါက နိုင်ငံရေး စကားစစ်ထိုးပွဲ (political debate) တစ်ခုကို တက်ရောက်လာသူတွေဆီကနေ debate မတိုင်ခင်နဲ့ ပြီးနောက် ကောက်ယူထားတဲ့ paired responses တွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီဟာ သီးခြား respondent တစ်ဦးစီကို ကိုယ်စားပြုပါတယ်။
- **debate_aggregate.sav.** ဒါက debate.sav ထဲက responses တွေကို aggregate လုပ်ထားတဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီဟာ debate မတိုင်ခင်နဲ့ ပြီးနောက် preference ရဲ့ cross-classification ကို ကိုယ်စားပြုပါတယ်။
- **demo.sav.** ဒါက လစဉ် offers တွေ ပေးပို့ဖို့ ရည်ရွယ်ချက်နဲ့ ဝယ်ယူထားတဲ့ customer database တစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Customer က offer ကို တုံ့ပြန်မှု (responded) ရှိမရှိ ဆိုတာကို demographic information အမျိုးမျိုးနဲ့အတူ မှတ်တမ်းတင်ထားပါတယ်။
- **demo_cs_1.sav.** ဒါက survey information တွေ ပါဝင်တဲ့ database တစ်ခု တည်ဆောက်ဖို့ ကုမ္ပဏီတစ်ခုရဲ့ ကြိုးပမ်းမှုတွေထဲက ပထမအဆင့်နဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက မတူညီတဲ့ မြို့တစ်မြို့ကို ကိုယ်စားပြုပြီး၊ region, province, district နဲ့ city identification တွေကို မှတ်တမ်းတင်ထားပါတယ်။
- **demo_cs_2.sav.** ဒါက survey information တွေ ပါဝင်တဲ့ database တစ်ခု တည်ဆောက်ဖို့ ကုမ္ပဏီတစ်ခုရဲ့ ကြိုးပမ်းမှုတွေထဲက ဒုတိယအဆင့်နဲ့ သက်ဆိုင်တဲ့ hypothetical data file

တစ်ခုပါ။ Case တစ်ခုစီက ပထမအဆင့်မှာ ရွေးချယ်ခဲ့တဲ့ မြို့တွေကနေ မတူညီတဲ့ household unit တစ်ခုကို ကိုယ်စားပြုပြီး၊ region, province, district, city, subdivision နဲ့ unit identification တွေကို မှတ်တမ်းတင်ထားပါတယ်။ ပထမအဆင့် နှစ်ခုက sampling information တွေကိုလည်း ထည့်သွင်းထားပါတယ်။

- **demo_cs.sav.** ဒါက complex sampling design ကို အသုံးပြုပြီး ကောက်ယူထားတဲ့ survey information တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက မတူညီတဲ့ household unit တစ်ခုကို ကိုယ်စားပြုပြီး၊ demographic နဲ့ sampling information မျိုးစုံကို မှတ်တမ်းတင်ထားပါတယ်။

- **diabetes_costs.sav.** ဒါက ဆီးချိုရောဂါ (diabetes) ရှိတဲ့ policy holders တွေနဲ့ ပတ်သက်ပြီး အာမခံကုမ္ပဏီတစ်ခုက ထိန်းသိမ်းထားတဲ့ သတင်းအချက်အလက်တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက မတူညီတဲ့ policy holder တစ်ဦးကို ကိုယ်စားပြုပါတယ်။

- **dietstudy.sav.** ဒီ hypothetical data file ထဲမှာ "Stillman diet" ဆိုတဲ့ လေ့လာမှုရဲ့ ရလဒ်တွေ ပါဝင်ပါတယ်။ Case တစ်ခုစီက သီးခြား subject တစ်ဦးကို ကိုယ်စားပြုပြီး၊ သူတို့ရဲ့ diet မတိုင်ခင်နဲ့ ပြီးနောက် ကိုယ်အလေးချိန်ကို ပေါင် (pounds) တွေနဲ့၊ triglyceride levels တွေကို mg/100 ml နဲ့ မှတ်တမ်းတင်ထားပါတယ်။

- **dmdata.sav.** ဒါက direct marketing ကုမ္ပဏီတစ်ခုအတွက် demographic နဲ့ purchasing information တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။ dmdata2.sav ထဲမှာ test mailing ရရှိခဲ့တဲ့ contacts တွေရဲ့ subset သတင်းအချက်အလက်တွေ ပါဝင်ပြီး၊ dmdata3.sav ထဲမှာတော့ test mailing မရရှိခဲ့တဲ့ ကျန်ရှိနေတဲ့ contacts တွေရဲ့ သတင်းအချက်အလက်တွေ ပါဝင်ပါတယ်။

- **dvdplayer.sav.** ဒါက DVD player အသစ်တစ်ခု တီထွင်ထုတ်လုပ်မှုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Prototype ကို အသုံးပြုပြီး marketing အဖွဲ့က focus group data တွေကို ကောက်ယူခဲ့ပါတယ်။ Case တစ်ခုစီက survey လုပ်ခံရတဲ့ သီးခြား user တစ်ဦးကို ကိုယ်စားပြုပြီး၊ သူတို့ရဲ့ demographic information တချို့နဲ့ prototype အပေါ် သူတို့ရဲ့ တုံ့ပြန်မှုတွေကို မှတ်တမ်းတင်ထားပါတယ်။

- **german_credit.sav.** ဒီ data file ကို University of California, Irvine ရဲ့ Repository of Machine Learning Databases ထဲက "German credit" dataset ကနေ ရယူထားတာပါ။

- **grocery_1month.sav.** ဒီ hypothetical data file က grocery_coupons.sav data file ထဲက weekly purchases တွေကို "rolled-up" လုပ်ထားတဲ့အတွက် Case တစ်ခုစီက သီးခြား customer တစ်ဦးကို ကိုယ်စားပြုပါတယ်။ အဲဒီရလဒ်အနေနဲ့ အပတ်စဉ် ပြောင်းလဲနေတဲ့ variables တချို့ ပျောက်ကွယ်သွားပြီး၊ အခု မှတ်တမ်းတင်ထားတဲ့ သုံးစွဲခဲ့တဲ့ ပမာဏ (amount spent) က လေ့လာမှုပြုလုပ်ခဲ့တဲ့ ၄ ပတ်တာ ကာလအတွင်း သုံးစွဲခဲ့တဲ့ ပမာဏတွေရဲ့ စုစုပေါင်း ဖြစ်ပါတယ်။

- **grocery_coupons.sav.** ဒါက သူတို့ customers တွေရဲ့ ဝယ်ယူမှု အလေ့အထကို စိတ်ဝင်စားတဲ့ ကုန်စုံဆိုင် ကွင်းဆက် (grocery store chain) တစ်ခုက ကောက်ယူထားတဲ့ survey data တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။ Customer တစ်ဦးစီကို ၄ ပတ်ကြာ စောင့်ကြည့်လေ့လာခဲ့ပြီး၊ case တစ်ခုစီက သီးခြား customer-week တစ်ခုကို ကိုယ်စားပြုကာ၊ အဲဒီအပတ်အတွင်း ကုန်စုံတွေအတွက် ဘယ်လောက်သုံးစွဲခဲ့တယ် ဆိုတာအပါအဝင် customer က ဘယ်နေရာမှာ ဘယ်လို ဈေးဝယ်တယ်ဆိုတဲ့ သတင်းအချက်အလက်တွေကို မှတ်တမ်းတင်ထားပါတယ်။

- **guttman.sav.** ဖြစ်နိုင်ခြေရှိတဲ့ လူမှုရေးအုပ်စု (social groups) တွေကို သရုပ်ဖော်ဖို့ Bell က ဇယားတစ်ခုကို တင်ပြခဲ့ပါတယ်။ Guttman က အဲဒီဇယားရဲ့ အစိတ်အပိုင်းတစ်ခုကို အသုံးပြုခဲ့ပြီး၊ အဲဒီထဲမှာ လူမှုရေးဆက်ဆံမှု (social interaction)၊ အုပ်စုတစ်ခုမှာ ပါဝင်ပတ်သက်နေတယ်လို့ ခံစားရမှု၊ အဖွဲ့ဝင်တွေရဲ့ ရုပ်ပိုင်းဆိုင်ရာ နီးကပ်မှု၊ နဲ့ ဆက်ဆံရေးရဲ့ တရားဝင်ဖြစ်မှု စတဲ့အရာတွေကို ဖော်ပြတဲ့ variables ၅ ခုကို၊ crowds, audiences, public, mobs, primary groups, secondary groups နဲ့ modern community အပါအဝင် သီအိုရီအရ social groups ၇ ခုနဲ့ ယှဉ်တွဲပြသထားပါတယ်။

- **health_funding.sav.** ဒါက health care funding (လူ ၁၀၀ လျှင် ပမာဏ)၊ disease rates (လူ ၁၀,၀၀၀ လျှင် နှုန်း) နဲ့ health care providers တွေဆီ လာရောက်ပြသမှု (လူ ၁၀,၀၀၀ လျှင် နှုန်း) တွေနဲ့ ပတ်သက်တဲ့ data တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက မတူညီတဲ့ မြို့တစ်မြို့ကို ကိုယ်စားပြုပါတယ်။

- **hivassay.sav.** ဒါက HIV infection ကို ရှာဖွေဖို့ rapid assay တစ်ခုကို တီထွင်ဖို့ ဆေးဝါးဓာတ်ခွဲခန်းတစ်ခုရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Assay ရဲ့ ရလဒ်တွေက အနီရောင် ရင့်သွားတဲ့ shades ၈ ခုဖြစ်ပြီး၊ အရောင်ပိုရင့်လေ infection ဖြစ်နိုင်ခြေ ပိုများလေလို့ ညွှန်ပြပါတယ်။ Blood samples ၂၀၀၀ ကို

ဓာတ်ခွဲခန်းစမ်းသပ်မှု လုပ်ဆောင်ခဲ့ရာမှာ၊ တစ်ဝက်က HIV infected ဖြစ်နေပြီး ကျန်တစ်ဝက်ကတော့ clean ဖြစ်နေပါတယ်။

- **hourlywagedata.sav.** ဒါက ရုံးနဲ့ ဆေးရုံရာထူးတွေမှာ အလုပ်လုပ်နေပြီး လုပ်သက် (experience) အဆင့်အမျိုးမျိုးရှိတဲ့ သူနာပြုတွေရဲ့ hourly wages တွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။

- **insurance_claims.sav.** ဒါက မသက်ာစရာကောင်းတဲ့၊ လိမ်လည်မှုဖြစ်နိုင်ချေရှိတဲ့ claims တွေကို သတိပေး (flagging) နိုင်မယ့် model တစ်ခု တည်ဆောက်ချင်တဲ့ အာမခံကုမ္ပဏီတစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက သီးခြား claim တစ်ခုကို ကိုယ်စားပြုပါတယ်။

- **insure.sav.** ဒါက client တစ်ဦးအနေနဲ့ 10-year term life insurance contract တစ်ခုအပေါ် claim လုပ်ဖို့ လိုအပ်မလားဆိုတာကို ညွှန်ပြတဲ့ risk factors တွေကို လေ့လာနေတဲ့ အာမခံကုမ္ပဏီတစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Data file ထဲက case တစ်ခုစီက age နဲ့ gender တူညီကြတဲ့ contracts တစ်စုံစီကို ကိုယ်စားပြုပြီး၊ တစ်ခုက claim ရှိတယ်လို့ မှတ်တမ်းတင်ထားကာ၊ နောက်တစ်ခုကတော့ မရှိပါဘူး။

- **judges.sav.** ဒါက ကျွမ်းကျင်တဲ့ ဒိုင်လူကြီးတွေ (ပေါင်း ဝါသနာရှင် တစ်ဦး) က ကျွမ်းဘားတင်ဆက်မှု (gymnastics performances) ၃၀၀ ကို ပေးထားတဲ့ scores တွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Row တစ်ခုစီက သီးခြား performance တစ်ခုကို ကိုယ်စားပြုပြီး၊ ဒိုင်တွေက တူညီတဲ့ performances တွေကိုပဲ ကြည့်ရှုခဲ့ကြတာပါ။

- **kinship_dat.sav.** Rosenberg နဲ့ Kim တို့က kinship terms ၁၅ ခုကို analyze လုပ်ဖို့ စတင်ခဲ့ပါတယ်။ သူတို့က ကောလိပ်ကျောင်းသား အုပ်စု ၄ စု (အမျိုးသမီး ၂ စု၊ အမျိုးသား ၂ စု) ကို အဲဒီ terms တွေကို ဆင်တူမှုအပေါ် အခြေခံပြီး sort လုပ်ခိုင်းခဲ့ပါတယ်။ အုပ်စု ၂ စု (အမျိုးသမီး ၁ စု၊ အမျိုးသား ၁ စု) ကို နှစ်ကြိမ် sort လုပ်ခိုင်းခဲ့ပြီး၊ ဒုတိယ sort က ပထမ sort နဲ့ မတူညီတဲ့ criterion ပေါ်မှာ အခြေခံထားတာပါ။ ဒါကြောင့် စုစုပေါင်း “sources” ၆ ခု ရရှိခဲ့ပါတယ်။ Source တစ်ခုစီက 15 x 15 proximity matrix တစ်ခုကို ကိုယ်စားပြုပြီး၊ အဲဒီ matrix ရဲ့ cells တွေဟာ source တစ်ခုမှာပါဝင်တဲ့ လူဦးရေထဲကနေ objects တွေကို အဲဒီ source မှာ အတူတကွ ခွဲခြားသတ်မှတ်ခဲ့တဲ့ (partitioned together) အကြိမ်အရေအတွက်ကို နုတ်ထားတဲ့ တန်ဖိုးနဲ့ ညီမျှပါတယ်။

- **kinship_ini.sav.** ဒီ data file ထဲမှာ kinship_dat.sav အတွက် three-dimensional solution တစ်ခုရဲ့ initial configuration ပါဝင်ပါတယ်။

- **kinship_var.sav.** ဒီ data file ထဲမှာ kinship_dat.sav ရဲ့ solution တစ်ခုက dimensions တွေကို interpret လုပ်ဖို့ အသုံးပြုနိုင်တဲ့ independent variables တွေဖြစ်တဲ့ gender, gener(ation) နဲ့ degree (of separation) တို့ ပါဝင်ပါတယ်။ တိတိကျကျပြောရရင်၊ အဲဒီ variables တွေကို solution ရဲ့ နေရာကို ဒီ variables တွေရဲ့ linear combination တစ်ခုဆီ ကန့်သတ်ထားဖို့ အသုံးပြုနိုင်ပါတယ်။

- **marketvalues.sav.** ဒီ data file က ၁၉၉၉-၂၀၀၀ ခုနှစ်တွေအတွင်း အီလီနွိုက်ပြည်နယ်၊ Algonquin ရှိ အိမ်ရာဖွံ့ဖြိုးမှုစီမံကိန်းအသစ်တစ်ခုက အိမ်အရောင်းအဝယ်တွေနဲ့ သက်ဆိုင်ပါတယ်။ ဒီအရောင်းအဝယ်တွေက public record ကိစ္စရပ်တွေပါ။

- **nhis2000_subset.sav.** National Health Interview Survey (NHIS) ဆိုတာ အမေရိကန် အရပ်သားလူဦးရေရဲ့ ကြီးမားတဲ့ population-based survey တစ်ခုပါ။ Interviews တွေကို nationally representative ဖြစ်တဲ့ households sample တွေထဲမှာ မျက်နှာချင်းဆိုင် ပြုလုပ်ခဲ့တာပါ။ Household တစ်ခုစီက အဖွဲ့ဝင်တွေအတွက် demographic information နဲ့ ကျန်းမာရေးဆိုင်ရာ အပြုအမူတွေ၊ အခြေအနေတွေအကြောင်း observations တွေကို ရယူထားပါတယ်။ ဒီ data file ထဲမှာ ၂၀၀၀ ပြည့်နှစ် survey ရဲ့ သတင်းအချက်အလက် subset တစ်ခု ပါဝင်ပါတယ်။

- **ozone.sav.** Data ထဲမှာ ကျန်ရှိနေတဲ့ variables တွေကနေ ozone concentration ကို predict လုပ်ဖို့အတွက် meteorological variables ၆ ခုအပေါ် ကောက်ယူထားတဲ့ observations ၃၃၀ ပါဝင်ပါတယ်။ ယခင် သုတေသီတွေက ဒီ variables တွေကြားမှာ nonlinearities တွေရှိနေတာကို တွေ့ရှိခဲ့ပြီး၊ ဒါက standard regression approaches တွေကို အဟန့်အတားဖြစ်စေပါတယ်။

- **pain_medication.sav.** ဒီ hypothetical data file ထဲမှာ နာတာရှည် အဆစ်အမြစ်ရောင်ရမ်းနာကျင်မှု (chronic arthritic pain) ကို ကုသဖို့အတွက် anti-inflammatory medication ရဲ့ clinical trial ရလဒ်တွေ ပါဝင်ပါတယ်။ အထူးစိတ်ဝင်စားစရာကောင်းတာက ဆေးစွမ်းပြဖို့ ကြာချိန်နဲ့ ရှိပြီးသား ဆေးတစ်မျိုးနဲ့ ယှဉ်ရင် ဘယ်လိုနေလဲ ဆိုတာပါပဲ။

- **patient_los.sav.** ဒီ hypothetical data file ထဲမှာ နှလုံးရောဂါ (myocardial infarction ဒါမှမဟုတ် MI) သံသယနဲ့ ဆေးရုံတက်ရတဲ့ လူနာတွေရဲ့ ကုသမှုမှတ်တမ်းတွေ ပါဝင်ပါတယ်။ Case တစ်ခုစီက သီးခြား လူနာတစ်ဦးကို ကိုယ်စားပြုပြီး သူတို့ဆေးရုံတက်စဉ် ကာလနဲ့ သက်ဆိုင်တဲ့ variables တွေကို မှတ်တမ်းတင်ထားပါတယ်။

- **patlos_sample.sav.** ဒီ hypothetical data file ထဲမှာ နှလုံးရောဂါ (MI) ကုသနေစဉ်အတွင်း သွေးခဲပျော်ဆေး (thrombolytics) ရရှိခဲ့တဲ့ လူနာတွေရဲ့ sample ကုသမှုမှတ်တမ်းတွေ ပါဝင်ပါတယ်။ Case တစ်ခုစီက သီးခြား လူနာတစ်ဦးကို ကိုယ်စားပြုပြီး သူတို့ဆေးရုံတက်စဉ် ကာလနဲ့ သက်ဆိုင်တဲ့ variables တွေကို မှတ်တမ်းတင်ထားပါတယ်။

- **poll_cs.sav.** ဒါက လွှတ်တော်ရွေးရောက်နေတဲ့ ဥပဒေကြမ်း (bill) တစ်ခုအပေါ် ပြည်သူတွေရဲ့ ထောက်ခံမှုအဆင့်ကို ဆုံးဖြတ်ဖို့ pollsters တွေရဲ့ ကြိုးပမ်းမှုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Cases တွေက မှတ်ပုံတင်ထားတဲ့ မဲဆန္ဒရှင်တွေကို ကိုယ်စားပြုပါတယ်။ Case တစ်ခုစီမှာ မဲဆန္ဒရှင် နေထိုင်တဲ့ county, township နဲ့ neighborhood တို့ကို မှတ်တမ်းတင်ထားပါတယ်။

- **poll_cs_sample.sav.** ဒီ hypothetical data file ထဲမှာ poll_cs.sav ထဲက မဲဆန္ဒရှင်တွေရဲ့ sample တစ်ခု ပါဝင်ပါတယ်။ Sample ကို poll.csplan plan file မှာ သတ်မှတ်ထားတဲ့ design အတိုင်း ကောက်ယူခဲ့တာဖြစ်ပြီး၊ ဒီ data file ထဲမှာ inclusion probabilities နဲ့ sample weights တွေကို မှတ်တမ်းတင်ထားပါတယ်။ သတိပြုရမှာက sampling plan က probability-proportional-to-size (PPS) method ကို အသုံးပြုထားတဲ့အတွက်၊ joint selection probabilities တွေပါဝင်တဲ့ ဖိုင်တစ်ခု (poll_jointprob.sav) လည်း ရှိပါတယ်။ မဲဆန္ဒရှင်တွေရဲ့ demographics နဲ့ အဆိုပြုထားတဲ့ ဥပဒေကြမ်းအပေါ် သူတို့ရဲ့ သဘောထားနဲ့ သက်ဆိုင်တဲ့ ထပ်ဆောင်း variables တွေကို sample ကောက်ယူပြီးနောက်မှ စုဆောင်းပြီး data file ထဲ ထည့်သွင်းထားတာပါ။

- **property_assess.sav.** ဒါက ကန့်သတ်ထားတဲ့ အရင်းအမြစ်တွေပေါ်မှာ အိမ်ခြံမြေတန်ဖိုး (property value) သတ်မှတ်ချက်တွေကို ခေတ်နဲ့အညီဖြစ်အောင် ထိန်းသိမ်းထားဖို့ county assessor တစ်ဦးရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Cases တွေက ပြီးခဲ့တဲ့နှစ်အတွင်း အဲဒီ county မှာ ရောင်းချခဲ့တဲ့ အိမ်ခြံမြေတွေကို ကိုယ်စားပြုပါတယ်။ Data file ထဲက case တစ်ခုစီမှာ အိမ်ခြံမြေတည်ရှိတဲ့ township၊ အဲဒီအိမ်ခြံမြေကို နောက်ဆုံးလာရောက်စစ်ဆေးခဲ့တဲ့ assessor၊ အဲဒီအကဲဖြတ်မှုကနေစပြီး ကြာချိန်၊ အဲဒီအချိန်က သတ်မှတ်ခဲ့တဲ့ တန်ဖိုးနဲ့ အိမ်ခြံမြေရဲ့ ရောင်းဈေး (sale value) တို့ကို မှတ်တမ်းတင်ထားပါတယ်။

- **property_assess_cs.sav.** ဒါက ကန့်သတ်ထားတဲ့ အရင်းအမြစ်တွေပေါ်မှာ အိမ်ခြံမြေတန်ဖိုး သတ်မှတ်ချက်တွေကို ခေတ်နဲ့အညီဖြစ်အောင် ထိန်းသိမ်းထားဖို့ state assessor တစ်ဦးရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Cases

တွေက အဲဒီ state ထဲမှာရှိတဲ့ အိမ်ခြံမြေတွေကို ကိုယ်စားပြုပါတယ်။ Data file ထဲက case တစ်ခုစီမှာ အိမ်ခြံမြေတည်ရှိတဲ့ county, township နဲ့ neighborhood၊ နောက်ဆုံးအကဲဖြတ်မှုကနေစပြီး ကြာချိန် နဲ့ အဲဒီအချိန်က သတ်မှတ်ခဲ့တဲ့ တန်ဖိုးတို့ကို မှတ်တမ်းတင်ထားပါတယ်။

- **property_assess_cs_sample.sav.** ဒီ hypothetical data file ထဲမှာ property_assess_cs.sav ထဲက အိမ်ခြံမြေတွေရဲ့ sample တစ်ခု ပါဝင်ပါတယ်။ Sample ကို property_assess.csplan plan file မှာ သတ်မှတ်ထားတဲ့ design အတိုင်း ကောက်ယူခဲ့တာဖြစ်ပြီး၊ ဒီ data file ထဲမှာ inclusion probabilities နဲ့ sample weights တွေကို မှတ်တမ်းတင်ထားပါတယ်။ ထပ်ဆောင်း variable ဖြစ်တဲ့ Current value ကို sample ကောက်ယူပြီးနောက်မှ စုဆောင်းပြီး data file ထဲ ထည့်သွင်းထားတာပါ။

- **recidivism.sav.** ဒါက အစိုးရ တရားဥပဒေစိုးမိုးရေး အေဂျင်စီတစ်ခုက သူတို့ရဲ့ စီရင်ပိုင်ခွင့်ဧရိယာအတွင်း ပြစ်မှုထပ်မံကျူးလွန်မှုနှုန်း (recidivism rates) ကို နားလည်ဖို့ ကြိုးပမ်းမှုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက ယခင်က ပြစ်မှုကျူးလွန်ခဲ့သူကို ကိုယ်စားပြုပြီး၊ သူတို့ရဲ့ demographic information၊ သူတို့ရဲ့ ပထမဆုံး ပြစ်မှုနဲ့ပတ်သက်တဲ့ အသေးစိတ်အချက်အလက် တချို့နဲ့၊ တကယ်လို့ ပထမအကြိမ် အဖမ်းခံရပြီး နှစ်နှစ်အတွင်းမှာ ဒုတိယအကြိမ် အဖမ်းခံရမယ်ဆိုရင် အဲဒီအချိန်အထိ ကြာချိန်တို့ကို မှတ်တမ်းတင်ထားပါတယ်။

- **recidivism_cs_sample.sav.** ဒါက အစိုးရ တရားဥပဒေစိုးမိုးရေး အေဂျင်စီတစ်ခုက သူတို့ရဲ့ စီရင်ပိုင်ခွင့်ဧရိယာအတွင်း ပြစ်မှုထပ်မံကျူးလွန်မှုနှုန်းကို နားလည်ဖို့ ကြိုးပမ်းမှုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက ၂၀၀၃ ခုနှစ်၊ ဇွန်လအတွင်း ပထမအကြိမ် အဖမ်းခံရမှုကနေ ပြန်လည်လွတ်မြောက်လာတဲ့ ယခင်က ပြစ်မှုကျူးလွန်ခဲ့သူကို ကိုယ်စားပြုပြီး၊ သူတို့ရဲ့ demographic information၊ သူတို့ရဲ့ ပထမဆုံး ပြစ်မှုနဲ့ပတ်သက်တဲ့ အသေးစိတ်အချက်အလက် တချို့နဲ့၊ တကယ်လို့ ၂၀၀၆ ခုနှစ်၊ ဇွန်လကုန်မတိုင်ခင် ဒုတိယအကြိမ် အဖမ်းခံရမယ်ဆိုရင် အဲဒီဒုတိယအကြိမ် အဖမ်းခံရတဲ့ရက်စွဲတို့ကို မှတ်တမ်းတင်ထားပါတယ်။ Offenders တွေကို recidivism_cs.csplan မှာ သတ်မှတ်ထားတဲ့ sampling plan အရ sampled departments တွေကနေ ရွေးချယ်ခဲ့တာပါ။ probability-proportional-to-size (PPS) method ကို အသုံးပြုထားတဲ့အတွက်၊ joint selection probabilities တွေပါဝင်တဲ့ ဖိုင်တစ်ခု (recidivism_cs_jointprob.sav) လည်း ရှိပါတယ်။

- **rfm_transactions.sav.** ဝယ်ယူတဲ့ရက်စွဲ (date of purchase)၊ ဝယ်ယူခဲ့တဲ့ ပစ္စည်း(များ) (item(s) purchased) နဲ့ transaction တစ်ခုစီရဲ့ ငွေကြေးပမာဏ (monetary amount) အပါအဝင် purchase transaction data တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။

- **salesperformance.sav.** ဒါက sales training courses အသစ်နှစ်ခုကို အကဲဖြတ်ခြင်းနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ ဝန်ထမ်း ၆၀ ကို အုပ်စု ၃ စု ခွဲထားပြီး အားလုံးက standard training ကို ရရှိကြပါတယ်။ အဲဒီအပြင် group 2 က technical training ကို ရရှိပြီး၊ group 3 ကတော့ hands-on tutorial ကို ရရှိပါတယ်။ သင်တန်းပြီးဆုံးချိန်မှာ ဝန်ထမ်းတစ်ဦးစီကို စစ်ဆေးခဲ့ပြီး သူတို့ရဲ့ score ကို မှတ်တမ်းတင်ခဲ့ပါတယ်။ Data file ထဲက case တစ်ခုစီက သီးခြား သင်တန်းသားတစ်ဦးစီကို ကိုယ်စားပြုပြီး၊ သူတို့ကို သတ်မှတ်ပေးခဲ့တဲ့ အုပ်စုနဲ့ စာမေးပွဲမှာ သူတို့ရရှိခဲ့တဲ့ score ကို မှတ်တမ်းတင်ထားပါတယ်။

- **satisf.sav.** ဒါက ဆိုင်ခွဲ (store locations) ၄ ခုမှာ retail ကုမ္ပဏီတစ်ခုက ပြုလုပ်ခဲ့တဲ့ satisfaction survey တစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Customers စုစုပေါင်း ၅၈၂ ဦးကို survey ပြုလုပ်ခဲ့ပြီး၊ case တစ်ခုစီက customer တစ်ဦးတည်းဆီက responses တွေကို ကိုယ်စားပြုပါတယ်။

- **screws.sav.** ဒီ data file ထဲမှာ screws, bolts, nuts, နဲ့ tacks တွေရဲ့ လက္ခဏာရပ်များနဲ့ ပတ်သက်တဲ့ သတင်းအချက်အလက်တွေ ပါဝင်ပါတယ်။

- **shampoo_ph.sav.** ဒါက ဆံပင်သုံးပစ္စည်းတွေ (hair products) အတွက် စက်ရုံတစ်ရုံရဲ့ quality control နဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ ပုံမှန် အချိန်အပိုင်းအခြားတွေမှာ သီးခြား output batches ၆ ခုကို တိုင်းတာပြီး သူတို့ရဲ့ pH ကို မှတ်တမ်းတင်ပါတယ်။ Target range က 4.5–5.5 ဖြစ်ပါတယ်။

- **ships.sav.** လှိုင်းတွေကြောင့် ကုန်တင်သင်္ဘော (cargo ships) တွေ ပျက်စီးမှုနဲ့ သက်ဆိုင်တဲ့ dataset တစ်ခုဖြစ်ပါတယ်။ ဖြစ်စဉ်အရေအတွက် (incident counts) တွေကို ship type, construction period, နဲ့ service period တို့အပေါ် မူတည်ပြီး Poisson rate အနေနဲ့ ဖြစ်ပွားတယ်လို့ model လုပ်နိုင်ပါတယ်။ Factors တွေရဲ့ cross-classification ကြောင့် ဖြစ်ပေါ်လာတဲ့ ဇယားရဲ့ cell တစ်ခုချင်းစီအတွက် aggregate months of service က risk နဲ့ ထိတွေ့မှုအတွက် (exposure to risk) တန်ဖိုးတွေကို ပေးစွမ်းပါတယ်။

- **site.sav.** ဒါက တိုးချဲ့လာတဲ့ စီးပွားရေးလုပ်ငန်းအတွက် နေရာအသစ်တွေ ရွေးချယ်ဖို့ ကုမ္ပဏီတစ်ခုရဲ့ ကြိုးပမ်းမှုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ သူတို့က

နေရာတွေကို သီးခြားစီ အကဲဖြတ်ဖို့ consultants နှစ်ဦးကို ငှားရမ်းခဲ့ပြီး၊ သူတို့က တိုးချဲ့အစီရင်ခံစာ (extended report) အပြင်၊ နေရာတစ်ခုစီကို အလားအလာရှိတဲ့နေရာအနေနဲ့ "good," "fair," ဒါမှမဟုတ် "poor" ဆိုပြီး အကျဉ်းချုံးပေးခဲ့ပါတယ်။

- **smokers.sav.** ဒီ data file ကို 1998 National Household Survey of Drug Abuse ကနေ ထုတ်နုတ်ထားတာဖြစ်ပြီး အမေရိကန် အိမ်ထောင်စုတွေရဲ့ probability sample တစ်ခု ဖြစ်ပါတယ်။ ဒါကြောင့် ဒီ data file ကို ခွဲခြမ်းစိတ်ဖြာဖို့အတွက် ပထမဆုံးအဆင့်အနေနဲ့ population trends တွေကို ထင်ဟပ်စေဖို့ data ကို weight ပေးရပါမယ်။

- **stocks.sav.** ဒီ hypothetical data file ထဲမှာ တစ်နှစ်စာအတွက် stocks prices တွေနဲ့ volume တွေ ပါဝင်ပါတယ်။

- **stroke_clean.sav.** ဒီ hypothetical data file ထဲမှာ Data Preparation option ထဲက procedures တွေကို သုံးပြီး ရှင်းလင်းထားပြီးသား (cleaned) medical database တစ်ခုရဲ့ အခြေအနေ ပါဝင်ပါတယ်။

- **stroke_invalid.sav.** ဒီ hypothetical data file ထဲမှာ medical database တစ်ခုရဲ့ ကနဦး အခြေအနေ ပါဝင်ပြီး၊ data entry errors တချို့လည်း ပါဝင်နေပါတယ်။

- **stroke_survival.sav.** ဒီ hypothetical data file က post-ischemic stroke ရဲ့ ပြန်လည်သန်စွမ်းရေး အစီအစဉ် (rehabilitation program) ကနေ ထွက်ခွာလာတဲ့ လူနာတွေရင်ဆိုင်ရတဲ့ စိန်ခေါ်မှုတွေနဲ့ survival times တွေအကြောင်းဖြစ်ပါတယ်။ လေဖြတ်ပြီးနောက်ပိုင်း (Post-stroke) မှာ myocardial infarction, ischemic stroke, ဒါမှမဟုတ် hemorrhagic stroke ဖြစ်ပွားမှုကို မှတ်သားထားပြီး ဖြစ်ပွားတဲ့အချိန်ကိုလည်း မှတ်တမ်းတင်ထားပါတယ်။ Sample ဟာ လေဖြတ်ပြီးနောက် ပြုလုပ်တဲ့ ပြန်လည်သန်စွမ်းရေး အစီအစဉ် ပြီးဆုံးတဲ့အထိ အသက်ရှင်ကျန်ရစ်သူ လူနာတွေကိုသာ ထည့်သွင်းထားတဲ့အတွက် left-truncated ဖြစ်နေပါတယ်။

- **stroke_valid.sav.** ဒီ hypothetical data file ထဲမှာ Validate Data procedure ကို သုံးပြီး values တွေကို စစ်ဆေးပြီးနောက် medical database တစ်ခုရဲ့ အခြေအနေ ပါဝင်ပါတယ်။ သူ့ထဲမှာ ပုံမှန်မဟုတ်တဲ့ (anomalous) cases တွေ ပါဝင်နေနိုင်ပါသေးတယ်။

- **survey_sample.sav.** ဒီ data file ထဲမှာ demographic data နဲ့ attitude measures အမျိုးမျိုးအပါအဝင် survey data တွေ ပါဝင်ပါတယ်။ သူ့ကို 1998 NORC General Social

Survey ထဲက variables တွေရဲ့ subset တစ်ခုအပေါ် အခြေခံထားတာဖြစ်ပေမယ့်၊ data values တချို့ကိုတော့ ပြင်ဆင်ထားပြီး ပြဿနိုင်ဖို့ရည်ရွယ်ချက်အတွက် ထပ်ဆောင်း ဇာတ်လမ်းဆင်ထားတဲ့ variables တွေ (fictitious variables) ကိုလည်း ထည့်သွင်းထားပါတယ်။

- **tcm_kpi.sav.** ဒါက စီးပွားရေးလုပ်ငန်းတစ်ခုအတွက် အပတ်စဉ် key performance indicators တွေရဲ့ values တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။ အဲဒီအချိန်ကာလအတွင်းမှာပဲ ထိန်းချုပ်လို့ရတဲ့ metrics တချို့အတွက် အပတ်စဉ် data တွေလည်း ပါဝင်ပါတယ်။

- **tcm_kpi_upd.sav.** ဒီ data file က tcm_kpi.sav နဲ့ အတူတူပဲဖြစ်ပေမယ့် နောက်ထပ် ၄ ပတ်စာအတွက် data တွေ ထပ်ပေါင်းပါဝင်ပါတယ်။

- **telco.sav.** ဒါက သူတို့ရဲ့ customer base ထဲမှာ churn (ဖောက်သည်လျော့ကျမှု) ကို လျှော့ချဖို့ telecommunications ကုမ္ပဏီတစ်ခုရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Case တစ်ခုစီက သီးခြား customer တစ်ဦးကို ကိုယ်စားပြုပြီး၊ demographic နဲ့ service အသုံးပြုမှု သတင်းအချက်အလက် အမျိုးမျိုးကို မှတ်တမ်းတင်ထားပါတယ်။

- **telco_extra.sav.** ဒီ data file က telco.sav data file နဲ့ ဆင်တူပေမယ့်၊ "tenure" နဲ့ log-transformed customer spending variables တွေကို ဖယ်ရှားလိုက်ပြီး standardized log-transformed customer spending variables တွေနဲ့ အစားထိုးထားပါတယ်။

- **telco_missing.sav.** ဒီ data file က telco.sav data file ရဲ့ subset တစ်ခုဖြစ်ပေမယ့်၊ demographic data values တချို့ကို missing values တွေနဲ့ အစားထိုးထားပါတယ်။

- **testmarket.sav.** ဒီ hypothetical data file က သူတို့ရဲ့ menu ထဲကို item အသစ်တစ်ခု ထည့်သွင်းဖို့ fast food chain တစ်ခုရဲ့ အစီအစဉ်တွေနဲ့ သက်ဆိုင်ပါတယ်။ Product အသစ်ကို ပရိုမိုးရှင်းလုပ်ဖို့ ဖြစ်နိုင်ခြေရှိတဲ့ campaigns သုံးခုရှိတဲ့အတွက်၊ item အသစ်ကို random အနေနဲ့ ရွေးချယ်ထားတဲ့ markets တွေက နေရာတွေမှာ မိတ်ဆက်ပေးခဲ့ပါတယ်။ နေရာတစ်ခုစီမှာ မတူညီတဲ့ promotion တစ်ခုကို အသုံးပြုပြီး၊ item အသစ်ရဲ့ weekly sales ကို ပထမ ၄ ပတ်အတွက် မှတ်တမ်းတင်ထားပါတယ်။ Case တစ်ခုစီက သီးခြား location-week တစ်ခုကို ကိုယ်စားပြုပါတယ်။

- **testmarket_1month.sav.** ဒီ hypothetical data file က testmarket.sav data file ထဲက weekly sales တွေကို "rolled-up" လုပ်ထားတဲ့အတွက် case တစ်ခုစီက သီးခြား location တစ်ခုကို ကိုယ်စားပြုပါတယ်။ အဲဒီရလဒ်အနေနဲ့ အပတ်စဉ် ပြောင်းလဲနေတဲ့ variables တချို့ ပျောက်ကွယ်သွားပြီး၊ အခု မှတ်တမ်းတင်ထားတဲ့ အရောင်း (sales) က လေ့လာမှုပြုလုပ်ခဲ့တဲ့ ၄ ပတ်တာ ကာလအတွင်း ရောင်းချခဲ့ရတဲ့ စုစုပေါင်း ဖြစ်ပါတယ်။
- **tree_car.sav.** ဒါက demographic နဲ့ vehicle purchase price data တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။
- **tree_credit.sav.** ဒါက demographic နဲ့ bank loan history data တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။
- **tree_missing_data.sav** ဒါက missing values တွေ အများကြီးပါဝင်တဲ့ demographic နဲ့ bank loan history data တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။
- **tree_score_car.sav.** ဒါက demographic နဲ့ vehicle purchase price data တွေ ပါဝင်တဲ့ hypothetical data file တစ်ခုပါ။
- **tree_textdata.sav.** Variables တွေကို measurement level နဲ့ value labels တွေ မသတ်မှတ်ရသေးခင်မှာ default အခြေအနေကို ပြသဖို့ အဓိကရည်ရွယ်ထားတဲ့၊ variables နှစ်ခုတည်းသာ ပါဝင်တဲ့ ရိုးရှင်းတဲ့ data file တစ်ခုပါ။
- **tv-survey.sav.** ဒါက အောင်မြင်နေတဲ့ အစီအစဉ်တစ်ခုရဲ့ ထုတ်လွှင့်မှုကို တိုးချဲ့ဖို့ စဉ်းစားနေတဲ့ TV studio တစ်ခုက ပြုလုပ်ခဲ့တဲ့ survey တစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ ဖြေဆိုသူ ၉၀၆ ဦးကို အခြေအနေအမျိုးမျိုးအောက်မှာ အဲဒီအစီအစဉ်ကို ကြည့်ဖြစ်မလားဆိုပြီး မေးမြန်းခဲ့ပါတယ်။ Row တစ်ခုစီက သီးခြား ဖြေဆိုသူတစ်ဦးကို ကိုယ်စားပြုပြီး၊ column တစ်ခုစီကတော့ သီးခြား condition တစ်ခုကို ကိုယ်စားပြုပါတယ်။
- **ulcer_recurrence.sav.** ဒီဖိုင်ထဲမှာ အစာအိမ်အနာ (ulcer) ပြန်ဖြစ်တာကို ကာကွယ်ဖို့ therapies နှစ်ခုရဲ့ ထိရောက်မှုကို နှိုင်းယှဉ်ဖို့ ရည်ရွယ်တဲ့ လေ့လာမှုတစ်ခုဆီကနေ ရရှိတဲ့ တစ်စိတ်တစ်ပိုင်း သတင်းအချက်အလက်တွေ ပါဝင်ပါတယ်။ ဒါက interval-censored data ရဲ့ ဥပမာကောင်းတစ်ခု ဖြစ်ပြီး၊ အခြားနေရာတွေမှာလည်း တင်ပြပြီး ခွဲခြမ်းစိတ်ဖြာ (analyzed) ခဲ့ပြီးဖြစ်ပါတယ်။
- **ulcer_recurrence_recoded.sav.** ဒီဖိုင်က လေ့လာမှုအဆုံးသတ်ရဲ့ event probability ကို ရိုးရိုးလေး မော်ဒယ်လုပ်မယ့်အစား လေ့လာမှုရဲ့ interval တစ်ခုစီအတွက် event probability

ကို model လုပ်လို့ရအောင် ulcer_recurrence.sav ထဲက သတင်းအချက်အလက်တွေကို ပြန်လည်စီစဉ် (reorganizes) ပေးပါတယ်။ ၎င်းကိုလည်း အခြားနေရာတွေမှာ တင်ပြပြီး ခွဲခြမ်းစိတ်ဖြာခဲ့ပြီးဖြစ်ပါတယ်။

- **verd1985.sav.** ဒီ data file က survey တစ်ခုနဲ့ သက်ဆိုင်ပါတယ်။ Subjects ၁၅ ဦးကနေ variables ၈ ခုအတွက် responses တွေကို မှတ်တမ်းတင်ထားတာပါ။ စိတ်ဝင်စားစရာ variables တွေကို sets ၃ ခု ခွဲထားပါတယ်။ Set 1 မှာ age နဲ့ marital ပါဝင်ပြီး၊ set 2 မှာ pet နဲ့ news ပါဝင်ကာ၊ set 3 မှာတော့ music နဲ့ live တို့ ပါဝင်ပါတယ်။ Pet ကို multiple nominal အနေနဲ့ scaled လုပ်ထားပြီး age ကိုတော့ ordinal အနေနဲ့ scaled လုပ်ထားပါတယ်။ ကျန်တဲ့ variables တွေအားလုံးကိုတော့ single nominal အနေနဲ့ scaled လုပ်ထားပါတယ်။

- **virus.sav.** ဒါက သူ့ရဲ့ networks တွေအပေါ် virus တစ်ခုရဲ့ အကျိုးသက်ရောက်မှုတွေကို ဆုံးဖြတ်ဖို့ Internet service provider (ISP) တစ်ခုရဲ့ ကြိုးပမ်းမှုတွေနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ သူတို့က networks တွေပေါ်က ရောဂါကူးစက်ခံထားရတဲ့ (infected) e-mail traffic ရဲ့ ခန့်မှန်းခြေ ရာခိုင်နှုန်းကို၊ အန္တရာယ်ကို စတင်တွေ့ရှိချိန်ကနေ ထိန်းချုပ်နိုင်တဲ့အချိန်အထိ အချိန်နဲ့အမျှ ခြေရာခံ (tracked) ခဲ့ကြပါတယ်။

- **wheeze_steubenville.sav.** ဒါက ကလေးတွေရဲ့ ကျန်းမာရေးအပေါ် လေထုညစ်ညမ်းမှုရဲ့ အကျိုးသက်ရောက်မှုတွေကို လေ့လာတဲ့ longitudinal study တစ်ခုဆီက subset တစ်ခုပါ။ Data ထဲမှာ အိုဟိုင်းယိုးပြည်နယ်၊ Steubenville က ကလေးတွေရဲ့ အသက် ၇၊ ၈၊ ၉ နဲ့ ၁၀ နှစ်အရွယ်တွေမှာ တိုင်းတာခဲ့တဲ့ wheezing status နဲ့ပတ်သက်တဲ့ repeated binary measures တွေ ပါဝင်ပြီး၊ လေ့လာမှုရဲ့ ပထမနှစ်အတွင်းမှာ မိခင်ဖြစ်သူက ဆေးလိပ်သောက်သူ ဟုတ်မဟုတ်ဆိုတဲ့ သတ်မှတ်ထားတဲ့ မှတ်တမ်း (fixed recording) လည်း ပါဝင်ပါတယ်။

- **workprog.sav.** ဒါက အခွင့်အလမ်းနည်းပါးသူတွေကို ပိုကောင်းတဲ့ အလုပ်အကိုင်တွေ ရရှိအောင် ကူညီပေးဖို့ ကြိုးစားနေတဲ့ အစိုးရရဲ့ အလုပ်အကိုင် အစီအစဉ် (works program) တစ်ခုနဲ့ သက်ဆိုင်တဲ့ hypothetical data file တစ်ခုပါ။ Program မှာ ပါဝင်နိုင်ခြေရှိတဲ့သူတွေရဲ့ sample တစ်ခုကို စောင့်ကြည့်လေ့လာခဲ့ပြီး၊ တချို့ကို program မှာ စာရင်းသွင်းဖို့ (enrollment) random အနေနဲ့ ရွေးချယ်ခဲ့ကာ တချို့ကိုတော့ မရွေးချယ်ခဲ့ပါဘူး။ Case တစ်ခုစီက သီးခြား program participant တစ်ဦးကို ကိုယ်စားပြုပါတယ်။

- **worldsales.sav** ဒီ hypothetical data file ထဲမှာ တိုက်ကြီးတွေ (continent) နဲ့ product အလိုက် sales revenue တွေ ပါဝင်ပါတယ်။